

UNIVERSITY OF ALBERTA
DYNAMIC MODELLING OF WASTEWATER TREATMENT PROCESSES

BY
AHMED GAMAL EL-DIN ©

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
ENVIRONMENTAL ENGINEERING

DEPARTMENT OF CIVIL & ENVIRONMENTAL ENGINEERING

EDMONTON, ALBERTA

FALL, 2002



National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-81192-1

Canada

UNIVERSITY OF ALBERTA
LIBRARY RELEASE FORM

NAME OF AUTHOR: AHMED GAMAL EL-DIN
TITLE OF THESIS: DYNAMIC MODELLING OF WASTEWATER
TREATMENT PROCESSES

DEGREE FOR WHICH
THIS THESIS
HAS PRESENTED: DOCTOR OF PHILOSOPHY

YEAR THIS DEGREE
GRANTED: 2002

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

(signed)-----



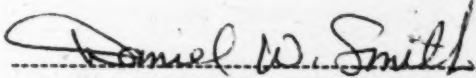
PERMANENT ADDRESS:
6716 Winnock Dr. Apt "A"
Indianpolis, Indiana
U.S.A
46220

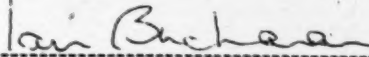
Date -----

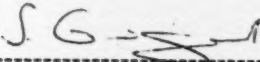
Aug 12, 2002

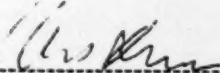
UNIVERSITY OF ALBERTA
FACULTY OF GRADUATE STUDIES AND RESEARCH

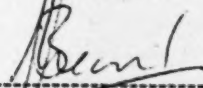
The undersigned certify that they have read, and recommended to the faculty of Graduate Studies and Research, for acceptance, a thesis entitled: DYNAMIC MODELLING OF WASTEWATER TREATMENT submitted by: Ahmed Gamal El-Din in partial fulfillment of the requirement for the degree of DOCTOR OF PHILOSOPHY IN ENVIRONMENTAL ENGINEERING.

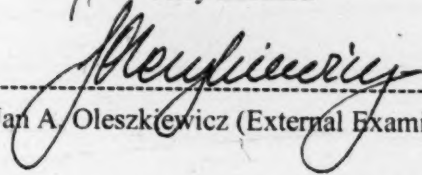

Daniel W. Smith-Supervisor


Ian D. Buchanan


Selma Guigard


Mohamed Al-Hussein


Jerry Leonard


Jan A. Oleszkiewicz (External Examiner)

Date Aug 7, 2002

TO

My Father Mr. Galal Gamal El-Din

My Mother Dr. Nadia El-Safari

My Wife Heba Salem

My Kids Yaseen and Malak

My Brother Dr. Mohamed Gamal El-Din

ABSTRACT

An artificial neural network (ANN) model was developed and used to provide short-term predictions of wastewater inflow rate that enters the Gold Bar Wastewater Treatment Plant (GBWWTP), the largest plant in the Edmonton area (Alberta, Canada). The neural network model uses rainfall data, observed in the collection system discharging to the plant, as inputs. The model was trained and validated using rainfall-flow data for the summers (May to October) of 1995 to 1997. The model has proven to be adequate for predicting the wastewater inflow rate.

Studying how and to what extent effluent total suspended solids (TSS) and chemical oxygen demand (COD) are related to influent TSS, COD, and flow in a full-scale primary sedimentation process was done using ANN and Box-Jenkins transfer function models. The analysis was based on data collected hourly over two periods of sampling, each lasted one week, conducted during the summer of 1999 at the GBWWTP. With the Box-Jenkins approach, stochastic and transfer function components were combined to form a dynamic model and the relative importance of the two components were quantitatively assessed. The models were proven to be adequate for predicting the TSS and COD concentrations in the primary effluent. The combined stochastic-transfer function models gave overall better predictions than the ANN models that utilized the same data set, however, the ANN models had the advantage that they only use information about the input data and no previous values of the output variable are used as inputs.

A hybrid model was used to predict the quality of effluent wastewater from a full-scale activated sludge process at the GBWWTP. Two components were included in the model: (1) a mechanistic component, which was a simplified version of the Activated Sludge Model No. 1 (ASM1) by the International Water Association (IWA); and (2) a black-box component that was an ANN in some cases while a time series component in other cases. The function of the black-box component was to predict the error between the actual experimental data and the predictions of the mechanistic model component. An extensive sampling campaign was conducted on one of the activated sludge tanks at the plant during the summer of 2000. By including the black-box component in the hybrid model the predictions were significantly improved.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude and appreciation to Dr. Daniel W. Smith for his guidance, support, continuous encouragement, and patience throughout my research program.

I also want to express my appreciation and thanks to my brother and friend Dr. Mohamed Gamal El-Din who has been always there when I needed him and with whom I shared great discussions throughout my research program. He also has provided me with great suggestions and thoughts throughout the preparation of my thesis.

I would to thank all the technical staff at the Environmental Engineering Laboratory at the University of Alberta for their great assistance during the laboratory program. Especially, Mr. Nick Chernuka, Mrs. Maria Demeter, and Mrs. Debra Long. I also appreciate the assistance by Mr. Garry Solonynko during the setting-up of the experimental apparatus.

The financial support provided by the City of Edmonton Drainage Branch is highly appreciated.

Finally a special thank you to my parents, my brother, and my wife Heba for their encouragement, understanding, and sacrifices during the course of the long journey of my graduate studies.

TABLE OF CONTENTS

CHAPTER 1.	GENERAL INTRODUCTION.....	1
1.1	BACKGROUND	1
1.2	DYNAMIC MODELS FOR WASTEWATER TREATMENT PROCESSES.....	3
1.2.1	Time Series Models	6
1.2.1.1	Application of Time Series Analysis in Environmental and Water Resources Engineering.....	7
1.2.2	Artificial Neural Networks (ANNs).....	8
1.2.2.1	Application of ANNs in Environmental Engineering.....	11
1.2.3	The Box-Jenkins Transfer Function Methodology vs. Artificial Neural Networks (ANNs)	15
1.3	STUDY AREA – THE GOLD BAR WASTEWATER TREATMENT PLANT (GBWWTP).....	16
1.4	PROBLEM DESCRIPTION.....	16
1.5	ORGANIZATION OF THIS DISSERTATION	17
1.6	REFERENCES	19
CHAPTER 2.	WASTEWATER INFLOW PREDICTION USING AN ARTIFICIAL NEURAL NETWORK	22
2.1	INTRODUCTION	22
2.2	DESCRIPTION OF THE AREA UNDER STUDY	23
2.3	OVERVIEW OF ANN	24
2.4	DEVELOPMENT OF THE MODEL.....	27
2.4.1	Input and Output Variables.....	27
2.4.2	Source Data Screening and Analysis	27
2.4.3	Architecture of the Neural Network Model.....	31
2.4.4	Training the Network.....	34
2.5	RESULTS AND DISCUSSION.....	36
2.6	CONCLUSIONS.....	51

2.7	REFERENCES	52
CHAPTER 3.	MODELLING A FULL-SCALE PRIMARY SEDIMENTATION TANK USING ARTIFICIAL NEURAL NETWORKS.....	53
3.1	INTRODUCTION	53
3.2	DYNAMIC MODELLING.....	55
3.3	DESCRIPTION OF THE TREATMENT PLANT.....	56
3.4	SAMPLING PROGRAM	58
3.5	METHODS	58
3.6	RESULTS OF THE SAMPLING PROGRAM	59
3.6.1	Total Suspended Solids Data	62
3.6.2	Chemical Oxygen Demand Data	65
3.7	MODELLING THE EFFLUENT TOTAL SUSPENDED SOLIDS AND CHEMICAL OXYGEN DEMAND.....	67
3.7.1	Theoretical Detention Time	67
3.7.2	Tracer Studies	68
3.7.3	Neural Network Models Developed	71
3.7.4	Structure of the Neural Network Models.....	72
3.7.5	Training the Networks	72
3.8	RESULTS OF THE NEURAL NETWORK MODELLING	73
3.8.1	The Neural Network Models Against Traditional Regression Models.....	80
3.9	DISCUSSION	80
3.10	CONCLUSIONS.....	81
3.11	REFERENCES	83
CHAPTER 4.	A COMBINED TRANSFER FUNCTION-NOISE MODEL TO PREDICT THE DYNAMIC BEHAVIOR OF A FULL-SCALE PRIMARY SEDIMENTATION TANK.....	85
4.1	INTRODUCTION	85
4.2	STUDY	87
4.3	QUALITATIVE DATA ANALYSIS.....	91

4.4	MODEL DEVELOPMENT	92
4.4.1	The Box-Jenkins Methodology	92
4.4.2	Model Identification and Estimation	95
4.4.3	Diagnostic Checking	97
4.4.3.1	Parameter Diagnostics	97
4.4.3.2	Residual Diagnostics	97
4.4.3.3	Closeness-of-Fit Statistics	106
4.5	QUANTITATIVE DATA ANALYSIS	106
4.5.1	Survey #1	108
4.5.2	Survey #2	110
4.5.3	Diagnostic Checking of the Models	111
4.5.4	Validating the Models	114
4.6	POSSIBLE APPLICATIONS OF THE MODELS	116
4.7	CONCLUSIONS	118
4.8	REFERENCES	120
CHAPTER 5.	DYNAMIC MODELLING OF A FULL-SCALE ACTIVATED SLUDGE PROCESS USING A SIMPLIFIED VERSION OF THE IWA-ASM1	122
5.1	INTRODUCTION	122
5.2	STUDY	124
5.2.1	The Gold Bar Wastewater Treatment Plant (GBWWTP)	124
5.2.2	The Activated Sludge Process at the GBWWTP	124
5.2.3	The Sampling Campaign	127
5.2.3.1	The First Survey – Plug Flow Mode	129
5.2.3.2	The Second Survey – Step-Feed Mode	131
5.2.3.3	Comparison between the Two Surveys	132
5.3	MODEL DEVELOPMENT	134
5.3.1	Conceptual Model	139
5.3.2	Components in ASM1 _{sp}	142
5.3.3	Processes in ASM1 _{sp}	144
5.4	STEADY STATE SIMULATIONS	149

5.4.1	Characterization of the Wastewater	151
5.4.2	Results – Plug Flow Mode	151
5.4.3	Results – Step-Feed Mode	156
5.5	DYNAMIC MODELLING UTILIZING THE EXPERIMENTAL DATA – RESULTS AND DISCUSSION	160
5.6	CONCLUSIONS	167
5.7	REFERENCES	168
CHAPTER 6.	A HYBRID MODEL FOR A FULL-SCALE ACTIVATED SLUDGE PROCESS THAT INCORPORATES A MECHANISTIC AND ANN/TIME SERIES COMPONENTS	170
6.1	INTRODUCTION	170
6.2	MODELLING THE ACTIVATED SLUDGE PROCESS	171
6.3	STUDY	173
6.3.1	The Sampling Campaign	174
6.4	DESCRIPTION OF THE HYBRID MODEL	180
6.4.1	The Mechanistic Component of the Hybrid Model; ASM1 _{SP}	181
6.4.2	The Black-Box Component of the Hybrid Model	182
6.5	MODEL DEVELOPMENT	185
6.5.1	ASM1 _{SP} - Unsteady State Simulations	185
6.5.2	The Hybrid Model	189
6.5.3	A Combined Transfer Function Noise Model for the Secondary Clarifier	193
6.6	RESULTS AND DISCUSSION	194
6.7	CONCLUSIONS	200
6.8	REFERENCES	202
CHAPTER 7.	CONCLUSIONS AND RECOMMENDATIONS	204
7.1	GENERAL OVERVIEW	204
7.2	CONCLUSIONS	206
7.3	RECOMMENDATIONS	210

LIST OF TABLES

Table 3.1	The physical parameters of primary settling tanks group 2 (PST 2).	57
Table 3.2	Descriptive statistics of the data.	60
Table 4.1	Data obtained from God Bar Wastewater Treatment Plant.	87
Table 4.2	Mathematical representation of the models.	107
Table 4.3	Model diagnostics.	112
Table 4.4	Ljung-Box white noise test for residuals.	113
Table 5.1	Physical dimensions of the activated sludge tanks.	125
Table 5.2	Average values for the parameters measured for the primary and secondary effluents.	132
Table 5.3	Average values for the parameters measured for the mixed liquor and return activated sludge.	133
Table 5.4	Process kinetics and stoichiometry for ASMSP.	135
Table 5.5	Different components in ASM1SP.	136
Table 5.6	Kinetic parameters in ASM1SP (Grady et al., 1999).....	138
Table 5.7	Stoichiometric coefficients in ASM1SP (Grady et al., 1999).....	139
Table 5.8	Wastewater characterization of the primary effluent.....	151
Table 5.9	Average primary effluent concentrations during the two sampling surveys.	161
Table 6.1.	The structure and parameter values for the time series models used as the error component in the hybrid model for predicting the soluble COD in the secondary effluent.....	191
Table 6.2.	The structure and parameter values for the models used to predict the total suspended solids concentration in the secondary effluent.	194

LIST OF FIGURES

Figure 2.1	The City of Edmonton drainage area	24
Figure 2.2	A schematic diagram of a multi-layer neural network.....	25
Figure 2.3	A histogram of the total amount of rainfall recorded during rainfall events.	28
Figure 2.4	Total inflow above dry weather flow.....	30
Figure 2.5	Maximum flow encountered at the entrance to the plant.....	30
Figure 2.6	Maximum flow encountered at the entrance to the plant.....	31
Figure 2.7	The number of epochs at which the maximum value of the R2 was obtained.....	37
Figure 2.8	The effect of increasing the size of moving window.....	38
Figure 2.9	Results obtained using a moving window of 8-hr.	39
Figure 2.10	Results obtained by the final model.....	40
Figure 2.11	Dry weather flow patterns simulated by the neural network model.	41
Figure 2.12	Actual versus predicted flow values for data set #1 – Part 1.	42
Figure 2.13	Actual versus predicted flow values for data set #1 – Part 2.	42
Figure 2.14	Actual versus predicted flow values for data set #1 – Part 3.	43
Figure 2.15	Actual versus predicted flow values for data set #1 – Part 4.	43
Figure 2.16	Actual versus predicted flow values for data set #1 – Part 5.	44
Figure 2.17	Actual versus predicted flow values for data set #1 – Part 6.	44
Figure 2.18	Actual versus predicted flow values for data set #1 – Part 7.	45
Figure 2.19	Actual versus predicted flow values for data set #1 – Part 8.	45
Figure 2.20	Actual versus predicted flow values for validation data set #2.	46
Figure 2.21	Actual versus predicted flow values for validation data set #3.	47
Figure 2.22	Prediction errors for the validation data.....	48
Figure 2.23	A portion of the training data set that contained erroneous flow data.	50
Figure 3.1	Wastewater flow recorded during the sampling periods.....	61
Figure 3.2	Dry weather flow profile.....	62
Figure 3.3	Influent total suspended solids.....	63

Figure 3.4	Influent total suspended solids profile during dry weather conditions - August 22 to 27, 1999.....	65
Figure 3.5	Influent chemical oxygen demand.....	66
Figure 3.6	Influent chemical oxygen demand profile during dry weather conditions - August 22 to 27, 1999.....	67
Figure 3.7	Theoretical detention time.....	68
Figure 3.8	Tracer studies.....	69
Figure 3.9	The effect of increasing the size of moving window.....	74
Figure 3.10	Results obtained by the TSS model.....	75
Figure 3.11	Predictions obtained by the TSS model for the training data set.....	76
Figure 3.12	Predictions obtained by the TSS model for the validation data set.....	77
Figure 3.13	Results obtained by the COD model.....	78
Figure 3.14	Predictions obtained by the COD model for the training data set.....	79
Figure 3.15	Predictions obtained by the COD model for the validation data set.....	79
Figure 4.1	Influent and effluent hourly data for the 7-days survey of June-July 1999: data set #1.....	88
Figure 4.2	Influent hourly flow and TSS, and effluent hourly TSS data for the 7-days survey of August 1999: data set #2.....	89
Figure 4.3	Influent and effluent hourly COD data for the 7-days survey of August 1999: data set #2.....	90
Figure 4.4	Cross correlation functions for prewhitened variables.....	96
Figure 4.5	Autocorrelation and partial autocorrelation functions for residuals.....	99
Figure 4.6	Cumulative periodogram check on residuals.....	100
Figure 4.7	Histogram of the residuals.....	102
Figure 4.8	Normal probability plot of the residuals.....	103
Figure 4.9	Residuals vs. predicted values.....	104
Figure 4.10	Residuals vs. input series.....	105
Figure 4.11	Outcome of the first tracer study conducted at high flow.....	109
Figure 4.12	One-step-ahead forecasts by model M-1.....	114
Figure 4.13	One-step-ahead forecasts by model M-2.....	115
Figure 4.14	One-step-ahead forecasts by model M-3.....	116

Figure 5.1	Activated sludge tanks #6, 7 and 8.	126
Figure 5.2	A schematic diagram of aeration tank #8 – Step-feed flow mode.	127
Figure 5.3	Wastewater flow measured at the effluent weir of each of the secondary clarifiers #6, 7 and 8 during the first survey.	129
Figure 5.4	COD of the primary effluent samples collected during the first survey.	130
Figure 5.5	TKN and ammonia nitrogen concentrations in the primary effluent samples collected during the first survey.	131
Figure 5.6	The traditional approach for modelling biomass decay (Adapted from Grady et al., 1999).	146
Figure 5.7	The death regeneration approach (Adapted from Grady et al., 1999).	147
Figure 5.8	A schematic of the configuration of tanks used to represent the activated sludge process.	149
Figure 5.9	Mixed liquor suspended solids (MLSS) concentration profiles – Plug flow mode.	152
Figure 5.10	Mixed liquor suspended solids (MLSS) concentration in CFSTR #8 – Plug flow mode.	152
Figure 5.11	Slowly biodegradable substrate concentration profiles – Plug flow mode.	153
Figure 5.12	Readily biodegradable substrate in secondary effluent – Plug flow mode.	154
Figure 5.13	Ammonia concentration in secondary effluent – Plug flow mode.	154
Figure 5.14	Nitrate concentration in secondary effluent – Plug flow mode.	155
Figure 5.15	Mixed liquor suspended solids (MLSS) concentration profiles – Step-feed mode.	156
Figure 5.16	Mixed liquor suspended solids (MLSS) concentration in CFSTR #8 – Step-feed mode.	157
Figure 5.17	Readily biodegradable substrate in secondary effluent – Step-feed mode.	158
Figure 5.18	Ammonia concentration in secondary effluent – Step-feed mode.	159

Figure 5.19	Nitrate concentration in secondary effluent – Step-feed mode.....	159
Figure 5.20	Ammonia nitrogen concentration in the secondary effluent – First survey.....	163
Figure 5.21	Nitrate nitrogen concentration in the secondary effluent – Second survey.....	164
Figure 5.22	Ammonia nitrogen concentration in the secondary effluent – Second survey.	165
Figure 5.23	Mixed liquor suspended solids concentration.....	166
Figure 6.1	Wastewater flow measured at the effluent weir of each of the secondary clarifiers #6, 7 and 8.	177
Figure 6.2	COD of the primary effluent samples.....	178
Figure 6.3	TKN and ammonia nitrogen concentrations in the primary effluent samples.....	179
Figure 6.4	The hybrid model.....	181
Figure 6.5	Correlogram of the residuals from the hybrid model that was used to predict the nitrate-nitrogen concentration in the secondary effluent for the second sampling survey.	185
Figure 6.6	A schematic of the configuration used to represent the equivalent activated sludge tank - Plug flow mode of operation.	187
Figure 6.7	Artificial neural network error-predictor for ammonia nitrogen concentration in secondary effluent.....	190
Figure 6.8	Monitoring the performance of the neural network error-predictor – Data of the first survey.....	191
Figure 6.9.	Correlograms of the autocorrelation and partial autocorrelation functions for the nitrate error series of the second survey.	192
Figure 6.10.	Actual data of the ammonia nitrogen concentration in the secondary effluent versus model predictions.....	195
Figure 6.11.	Actual data of the nitrate nitrogen concentration in the secondary effluent versus model predictions.	196
Figure 6.12.	Actual data of the soluble COD concentration in the secondary effluent versus model predictions.	197

Figure 6.13. Actual data of the total suspended solids concentration in the
secondary effluent versus model predictions.....198

CHAPTER 1. GENERAL INTRODUCTION

1.1 BACKGROUND

Since wastewater treatment consists of a sequence of complex physical and biological processes, currently in most wastewater treatment plants (WWTPs), process control is generally accomplished through operator's experience. Due to the lack of reliable instrumentation and the harsh environment in which these instrumentation are housed and operated, the introduction of advanced control technology in WWTPs has been slow. However, within the past few years, the topic of real time control (RTC) of WWTPs has quickly become a promising and active area of research and the improvement of the computer-based infrastructure for WWTPs has been identified as an important priority. Many factors have contributed to this surge of interest in the field of WWTP automation and control. Among those factors are: (1) increasingly stringent legislation concerning pollutant discharge; (2) increased concern about the effects that untreated or partially treated sewage has on surface water quality; (3) advances and investment in communication, computing, and sensor technology; and (4) the potential for long-term cost-effectiveness. Thus, providing autonomous, reliable and stable process control with highly efficient throughput at minimum cost has been recently identified as the long-term objective of wastewater treatment process operation (Katebi et al., 1999). Novotny et al. (1992) identified the alternative objectives, that have been presented in the literature, of RTC of urban sewerage systems to be: (1) minimization of discharges of untreated overflows; (2) stability of the treatment process and compliance with effluent standards; (3) minimization of total pollution load; and (4) avoidance or minimization of "bottleneck" situations, in which the hydraulic and/or mass loading capacity of a process unit within the system is exceeded.

In existing wastewater treatment plants, current control strategies are highly site-specific and depend on the level of technology available at the plant. Many older plants possess little or no monitoring equipment; such plants usually rely heavily on the experience of the operator. Control strategies are usually quite simple, e.g., maintaining a

constant mixed liquor suspended solids (MLSS), sludge residence time (SRT), or food to microorganism ratio (F/M), based on simple analytical measurements or observations (Gall and Patry, 1989). Many newer plants are equipped with a wide range of monitoring equipment, including dissolved oxygen (DO), suspended solids (SS), pH, and/or a turbidity meter, for real time monitoring and control. In such plants, operational strategies are usually based on traditional control algorithms applied to local control situations, e.g., air flow control to maintain a preset dissolved oxygen level in the aeration basin. In the past few years more research has been directed towards the field of dynamic modelling of wastewater processes for the purpose of real-time control. By far, the activated sludge process has gained the most attention. Enhanced primary sedimentation with the supplemental addition of chemicals and anaerobic digestion are examples of other unit processes, on which few studies exist in the literature that describe using a RTC system to optimize the performance of the process (Wilcox et al., 1995; Premier et al., 1999).

The main components that compromise a RTC system are: (1) a measurement device (sensor) that is used to monitor the process, for example, a turbidity meter; (2) a corrective regulator (actuator) that manipulates the process, for example, gates, valves, and similar devices; (3) a controller that causes the regulator to bring the process back to its desirable value (set point); and (4) a communication system by which the measured data are carried from the sensor to the controller and the signals of the controller are carried to the regulator. The controller of a RTC system uses a dynamic model of some kind in order to relate the inputs that it receives from the measurement device (or devices) to an output that is sent to the regulator in order to correct the process.

The capacity to control successfully the dynamic behavior of a system depends essentially on three factors (Beck, 1989): (1) the ability to observe the state of the process and its response to various perturbations (i.e., monitoring); (2) the ability to relate inputs to outputs (causes to effects) through the interpretation of field observations; and (3) the capacity to act, i.e., to manipulate the controlled inputs in order to control undesirable effects or to bring about more desirable effects. In order to obtain an effective RTC of a WWTP, the following dynamic processes have to be modelled: (1) input to the system,

i.e., rainfall-runoff-sewer flow and quality models; and (2) system response to the input, i.e., models that have the ability to relate causes (inputs, controls) to effects (outputs, responses).

In order to relate the inputs of an unsteady state system to its output, dynamic modelling techniques have to be utilized. In the terminology of process control, building a dynamic model for a system is referred to by "system identification". System identification can be said to be about the use of models for the interpretation of field data (Beck, 1989). The procedure of system identification comprises the following component parts: (1) experimental design; (2) selection of the type of model and the model structure; (3) estimating the values of the model parameters (i.e. calibrating the model); and (4) verification of the calibrated model. The concepts of model identifiability and experimental design are closely intertwined (Beck, 1989). An unidentifiable model structure can be made identifiable by an appropriate choice of the input and output vectors. Selecting the type of model used to represent a dynamic system for the purpose of real time process control is very important. A model for control purposes should not be too complex (Olsson, 1989). The many interacting species and substrates in a wastewater treatment plant, however, suggest extremely complex models. The resolution of the trade-off between relatively simple on-line models and complex-structured models seems to be the major problem in automation of biological treatment plants (Olsson, 1989).

1.2 DYNAMIC MODELS FOR WASTEWATER TREATMENT PROCESSES

Dynamic models of wastewater treatment processes can be divided into two main categories; linguistic and mathematical models. Linguistic models, such as expert systems, can relate cause to effect and they do not demand the construction of a mathematical model. Linguistic models are most suitable for describing phenomena in environmental systems that are very hard to represent by a mathematical model. The dynamics of algal populations, the clarification and thickening functions of the secondary clarifier, or the development of a bulking sludge are likely candidates (Beck, 1989). Most of the expert systems that have been developed in the field of wastewater treatment have

been directed into diagnosis. This category of applications helps identify the causes of and remedies for malfunctions in pollution control facilities. Diagnostic expert systems have been developed for activated sludge wastewater treatment systems and anaerobic digesters (Rossman and Siller, 1987; Rossman, 1989; Bruce and Party, 1989).

Mathematical models can take many forms. At one extreme, some models are highly mechanistic. These are most useful to the researcher seeking to understand the events occurring in a system. Generally such models are deterministic and incorporate direct links between inputs and outputs through ordinary and partial differential equations that seek to mimic reactions mechanisms. Such models are large in scope and require a detailed knowledge of the system and the evaluation of a large number of parameters. An example is the Activated Sludge Model No. 2 (ASM2) developed in 1995 by the International Water Association (IWA; formerly the International Association on Water Quality IAWQ) task group which is an extension of the very well known Activated Sludge Model No. 1 (ASM1). The full version of ASM2 contains 19 processes, 19 variables, 43 kinetic parameters, and 22 stoichiometric parameters.

Deterministic models can be divided into two categories; lumped-parameter models and distributed-parameter models (Beck, 1989). Lumped-parameter models assume that all the properties of the system are uniform across a given spatial volume of the reactor. These models imply complete mixing of the fluids within the designated volume, and frequently, this volume is taken to be the entire volume of the reactor. Therefore, a lumped-parameter model is a continuous flow stirred tank reactor (CFSTR) approximation. The lumped-parameter model has been the primary vehicle for the development of chemical, biochemical, and microbiological hypotheses about the behavior of unit processes of wastewater treatment (Beck, 1989). In a lumped-parameter model, a vector of state variables is used to quantify the internal state of the process (usually in terms of concentrations of dissolved or suspended substances) within the designated volume (space), and hence, the lumped-parameter model has its alternative designation as a state-space model. Because the majority of control problems focus on the

manipulation of changes with respect to time, they are strongly geared to the use of lumped-parameter models (Beck, 1989).

All systems exhibit behavior that varies with both time and space, and for some processes, both spatial and temporal variability may be important to the behavior of the process in practice, for example, a predominantly plug-flow reactor (PFR). In such situations, a distributed-parameter model is used to describe the system, in which, state variables are allowed to vary in time and space. The obvious motivation for the development of a distributed-parameter model is that fluid and mass transport phenomena are crucial to the dynamic behavior of a given process, and in such cases, momentum and energy balances may form part of the model, with fluid velocity being a constituent element of the state vector (Beck, 1989). Instead of expressing spatial variation explicitly in the model, as in the case of a distributed-parameter model, a system can be simulated by a series of CFSTR elements, in which case spatial variations would be implicit in the model.

Much more research will be required before it will be possible to develop mechanistic models which accurately describe the interactions between various organic substances, both inhibitory and biogenic, in wastewater treatment systems. Therefore, another type of model has been used by researchers in order to represent hard-to-describe systems. These are called empirical or "black-box" models. They are most useful to the operator because they can reflect real world responses. They are, however, highly system specific. As a result, they have to be developed specifically for the set of data under study, and hence, they are not easily transportable to new situations. Examples of the class of black box models are time series models and Artificial Neural Networks (ANNs).

Between the extremes of highly mechanistic models and black-box models lie many alternatives. A dynamic model may contain a mechanistic component and a black-box component. In such a hybrid model, the objective of the mechanistic component is to describe the physical phenomena within the system that is thought to be well-understood. The role of the black-box component is to bias the predictions by the mechanistic

component towards the experimental data. As a result, the black-box component of the hybrid model will compensate for the processes that were not included as part of the mechanistic model component because knowledge about them is not yet sufficient to allow this to be done. As the understanding of the processes that were not included in the mechanistic model because of lack of knowledge improves, they can be added to the model resulting in reduced load on the black-box (or error-predictor) component of the hybrid model.

1.2.1 Time Series Models

Time series models can be used in three important areas of application: forecasting of future values of a time series from current and past values; the determination of a transfer function of a system, i.e., the determination of a dynamic input-output model that can show the effect on the output of a system subject to inertia, of any given series of inputs; and the design of simple feed-forward and feedback control schemes by means of which potential deviations of the system output from a desired target may be compensated, so far as possible. Forecasting time series can provide a basis for control and optimization of industrial processes (Box and Jenkins, 1976).

Time series forecasting methods are analytical in nature; i.e., they use a variety of mathematical and statistical concepts and techniques to extract information from time series data, establish relationships among relevant factors, and extrapolate past behavior into the future. A wide class of processes called autoregressive-integrated moving-average processes "ARIMA" (often called "Box-Jenkins" methodology) provides a range of models that adequately represent many time series found in practice. The types of models employed in the Box-Jenkins methodology have been around for many years, however, it was not until the 1960's that Box and Jenkins recognized the importance of these models in the area of economic forecasting and developed the Box-Jenkins forecasting methodology to take advantage of them (Hoff, 1983).

The important characteristic of these models is that they were developed in terms of statistical concepts under the assumption that the processes being modelled are dynamic and subject to statistical fluctuations. The assumption is quite compatible with the behavior of environmental time series. The Box-Jenkins method is often referred to as univariate Box-Jenkins. The methodology and concepts employed in univariate Box-Jenkins, however, have also been extended to the cause-and-effect approach. The resulting method is known as multivariate Box-Jenkins or the Box-Jenkins transfer function method. The study of process dynamics is a topic of considerable industrial interest and the reason for that is to achieve better control of existing plants and to improve the design of new plants. Several methods have been proposed for estimating the transfer function of plant units from process records. Classical methods for estimating transfer function models based on deterministic perturbations of the input, such as step, pulse, and sinusoidal changes, have not always been successful, and this because the response of the system may be masked by uncontrollable disturbances collectively referred to as "noise" (Box and Jenkins, 1976). Statistical methods for estimating transfer function models that make allowance for noise, such as the Box-Jenkins transfer function models, can be employed with adequate results for the identification of dynamic systems.

1.2.1.1 Application of Time Series Analysis in Environmental and Water Resources Engineering

Huck and Farquhar (1974) applied ARIMA modelling techniques to model both dissolved oxygen and chloride concentrations in the St. Clair River near Corcenna, Ontario. The Box-Jenkins methodology was successful in modelling hourly water quality data by explaining 60% to 70% of the variation while producing adequate forecasts. Berthouex et al. (1975) utilized a first order auto-regressive model to study hourly influent biochemical oxygen demand (BOD) series data collected at the Madison Nine Springs plant, Wisconsin, USA. They concluded that time series models derived from actual field data show promise for describing system performance. Barnes and Rowe (1978) applied time series analysis to model volumetric flow data from the combined sewer treatment plants in Springfield, Ohio, USA and in Anderson, Indiana, USA. Chow

et al. (1983) utilized transfer function models in order to predict flood levels of the Saint John River at Fredericton, New Brunswick. The stochastic dynamic models produced reasonable forecasts for lead times up to two days. Caroni et al. (1984) successfully utilized single-input single-output transfer function models for determination of the rainfall-runoff-sediment yield relationship in the Pigeon Roost Creek in Mississippi, USA. Gurnell and Fenn (1984) studied the use of Box-Jenkins transfer function models applied to suspended sediment concentration-discharge relationships in a stream. Of all models considered, the transfer function gave by far the best forecasts. Thomstone et al. (1985) successfully demonstrated that starting with a "de-seasonalized" ARMA model of the inflow series and adding transfer functions for the rainfall and snowmelt series could provide a useful model in predicting quarter-monthly (i.e., near-weekly) natural inflows to the St. Jean reservoir in the province of Quebec, Canada. Lee (1986) studied the application of time series analysis to treatment plant data including BOD, water supply, and influent series. It was found that the Box-Jenkins type ARIMA models can be successfully used to model a treatment plant system. Jayawardena and Lai (1989) studied the mean monthly water quality data in the Pearl River in southern China using Box-Jenkins type ARIMA model. It was found that the forecasting of future water quality data using a Box-Jenkins type model was satisfactory. Capodaglio et al. (1990) applied an on-line system-identification technique, including a recursive estimation algorithm, to obtain one-step-ahead predictions of the daily influent inflow to the Fusina Wastewater Treatment Plant in Italy utilizing daily rainfall data. Lemke (1991) used single input-single output and multiple input-single output transfer function models for daily suspended sediment concentration for two drainage basins in Iowa.

1.2.2 Artificial Neural Networks (ANNs)

Artificial neural networks go by many names such as connectionist models and parallel distributed processing models. The ANN modelling approach is an artificial intelligence technique that attempts to simulate some important features of the human nervous system; in other words, the ability to solve problems by applying information gained from past experience to new problems or case scenarios. Artificial neural network

ideas were developed in the fifties, went into decline in the sixties, and only in the middle eighties did their true potential become evident. A book credited with the renewed interest was Rumelhart and McClelland's *Parallel Distributed Processing* (Rumelhart et al., 1986), which was published in 1987 and contains, among other things, the first well-known description of the back-propagation training algorithm. Analogous to a human brain, an ANN uses many simple computational elements, named artificial neurons, connected by variable weights. Although each neuron, alone, can only perform simple computations, the hierarchical organization of a network of interconnected neurons, or architecture, makes an ANN capable of performing complex tasks such as pattern classification and prediction.

Among the many different architectures, the multi-layer perceptron architecture is commonly used for prediction, in which neurons are arranged in layers, an input layer, hidden layers, and an output layer. Input neurons, or nodes, receive values of an instance of the input parameters that are fed to the network after being scaled into a numeric range that is efficient for calculations by the neural network model. Outcomes to the output parameters for the instance under consideration are assigned by the output neurons. Hidden neurons connect the input neurons to the output neurons and provide nonlinearity to the network. Each neuron is connected to every neuron in adjacent layers by a connection weight, which determines the strength of the relationship between two connected neurons. The output from a neuron is multiplied by the connection weight before being introduced as input to the neuron in the next layer. Each neuron, except those in the input layer, sums all of the inputs that it receives and the sum is converted to an output value according to a predefined activation, or transfer, function.

For prediction problems, a supervised learning algorithm is often utilized for teaching the network how to relate input data patterns to output data. The main reason for the popularity of the multi-layer perceptron architecture in recent years is the development of new training algorithms like the back-propagation algorithm. The algorithm requires a continuous differentiable nonlinearity to be used as the transfer function by the neurons. The back-propagation algorithm consists of the following steps:

(1) all weights are assigned initial small random values; (2) input data are propagated in a feed-forward manner through the network to produce output data according to the weights and the transfer function; (3) the outputs produced are compared with the target outputs which are known in advance; (4) the errors generated are then propagated backwards in a certain manner through the network for adjustments of the present weights using two factors, namely, a learning factor and a momentum factor; and (5) this procedure is repeated for each training example in the training set; a cycle (an epoch) represents one pass over the whole training set; multiple epochs are required until a satisfactory data mapping is achieved.

The main advantage of neural network models over a parametric approach is that neural networks can generate complicated relationships through examination of only the data points in the training set without assuming a pre-specified functional form. The prediction capability of neural networks, and the ability to quickly adapt to the changes in the status of a dynamic process make them easily-integrated into a control scheme. Although the neural network approach has its advantages, it also suffers from a few limitations. The traditional limits on learning speed and network size have been reduced by hardware improvement and more efficient algorithms. The main limitation that still exists is the one resulting from insufficient database for training the neural network. Two cases may occur: (1) an important input that is not included in the model or not measurable; and (2) scarcity of learning points in a certain region, which may cause the ANN to extrapolate to erroneous results without warning. To avoid the first case, the modelled physical system has to be well-represented by a set of measurable input and output variables. The remedy to the second case is trying to collect a data set that includes all possible scenarios for the system under study. Even so, a neural network may perform successfully at the beginning but after a while, its validity starts to decrease because of changes in the modelled system or operating conditions that occurred after training and have not been seen before by the neural network. As a result, a regular training schedule must be established as a check of the neural network validity.

1.2.2.1 Application of ANNs in Environmental Engineering

Capodaglio et al. (1991) applied artificial neural network models to the analysis of bulking conditions. The sludge volume index (SVI) was used to represent bulking conditions. The artificial neural system analysis was conducted using a time series input scheme. A feed-forward neural network was utilized. The input to the network consisted of time-lagged values of BOD/N ratios, N/P ratios, mixed liquor DO, mixed liquor temperature, and the food to microorganism ratio. A 5-day lag was chosen as it corresponded to the average mean cell residence time of the system. Each input pattern fed to the system consisted of five values representing the previous 5 days for each of the input variables and the one-day-ahead prediction of SVI value was the output of the network. The back-propagation algorithm was used in training. It was found that the neural network has a better prediction capability than the other stochastic models investigated.

Su and McAvoy (1992) studied a pilot scale wastewater treatment facility with biological removal of C, N, and P. Three networks were used to model the three processing units, the anaerobic pretreatment column, the aeration tanks and the sedimentation vessel, individually. Aeration was intermittent to achieve denitrification. For the course of 4 days, samples were taken every 7 minutes from the influent and effluent of each of the three units and were analyzed for NH_4^+ , NO_3^- , and PO_4^{3-} . The oxygen consumption rates at the aeration tanks were also recorded. The outputs of interest were the concentration measurements for the effluent of the three processing units while the inputs were the concentration measurements for the influent and the oxygen consumption rates. The one-step-ahead prediction by a feed-forward network for each system was found to be good.

Tyagi and Du (1992) demonstrated the application of neural network techniques for kinetic model building of heavy metals inhibition in activated sludge process. Feed-forward neural network based models for the cases of both unacclimated and acclimated microorganisms were developed. After trying different activation functions and different

sizes of the hidden layer, the sigmoid function was chosen and 6 neurons were used. The learning procedure of the neural network was based on the back-propagation algorithm. The task of the neural network model was to predict the maximum specific growth rate corresponding to different combinations of heavy metals present in the wastewater. The ratio of the maximum specific growth rate in the presence and that in the absence of metal, μ_m/μ_m^0 , as the index for the growth rate was the model output, while the ratio of the metal concentration and the initial concentration of the suspended solids, M/X_0 , was the model input data. The ANN model was able to predict the index of the growth rate μ_m/μ_m^0 .

A neural network approach was proposed by Tabak and Govind (1993) in trying to estimate a biodegradation kinetic constant of a compound based on the type and number of chemical groups that compose this compound. After a linear first order model was tested, it was found that it was necessary to develop a nonlinear method in order to incorporate the nonlinear interactions between the chemical groups. A three layer neural network was used with eight input nodes, eight hidden nodes, and one output node. The sigmoid transfer function was used in the hidden layer. Each input node corresponded to a specific functional group. The inputs to the network corresponded to the number of each type of chemical group present in the chemical structure. The biodegradation kinetic constant value was the output of the network. The back propagation algorithm was used in the training of the network. Results showed that the method was able to predict the kinetic constant values for compounds that were not in the training set.

A simple ANN model that takes into account the bioreactor, the settler, and their interaction, was built by Tyagi et al. (1993) for the prediction of the recycle ratio and the wastage rate that are required to have effluent substrate concentrations at a given set-point using the influent flow rate and influent substrate concentration as inputs to the model. A steady-state simulation using mass balance equations around the bioreactor and the limiting flux theory for the sedimentation analysis was used to generate the data patterns for training and testing the network. The sigmoid function was used as the transfer function and the back-propagation algorithm was used in the training of the

network. The hidden layer comprised of only two neurons. It was found that the neural network model exhibited an excellent performance in terms of both data fitting and prediction for the system. Trying to fit the same data, different polynomial equations with different orders were tested, however, the performance of the ANN model was still much better.

Tyagi et al. (1993) collected data from a pilot plant and analyzed it using a feed-forward ANN model to study the effect of the recycle ratio, α , and the clarifier overflow rate, U_o , on the concentration of the mixed liquor suspended solids, X , entering into the settling tank and the recycle suspended solids, X_r , leaving from the bottom of the tank. The data showed that the relation between those variables has a certain degree of nonlinearity. A neural network was used to predict the values of X and X_r based on the values of α and U_o at steady state condition. It was found that an adequate agreement between prediction and observation was achieved.

A neural-net-based model was developed by Du et al. (1994) to predict the solubilization of six heavy metals from sewage sludge using the bioleaching process. The sigmoid function was used as the transfer function. One hidden layer was used and it comprised of 16 nodes. Data from batch reactor experiments were used in training the network using the back-propagation algorithm. It was shown that the network is capable of predicting the metal solubilization using the pH and the initial concentrations of metals in the sludge as inputs to the network. The ANN model was used to predict the pH (which can be easily monitored during the course of the bioleaching process) at which the bioleaching process should be terminated which is a major advantage in terms of process control (Du et al., 1994).

An ANN was used by Pu and Hung (1994) to predict the performance of a medium-sized municipal wastewater treatment plant using rotating biological contactors and activated sludge process in treating medium strength municipal wastewater. Two years of average daily data were used in training and testing the network. The data used in the ANN modelling included raw wastewater flow rate, influent and effluent TSS and BOD₅

of primary settling tank and secondary settling tank. A multi-layer network was trained using the back propagation algorithm. One hidden layer was used and the number of nodes in the hidden layer was varied from 1 to 8. The sigmoid function was used as the transfer function in the hidden layer. Two general ANN models were used to predict the five-days biochemical oxygen demand (BOD₅) and the suspended solids (SS) concentrations in the secondary effluent. To test the validity of ANN model, the ANN results were compared with those of a multiple regression analysis. The results of the ANN models were superior.

A procedure has been developed by Cote' et al. (1995) to improve the accuracy of an existing mechanistic model of the activated sludge process, previously described by Lessard and Beck (1993). The first step was the optimization of the numerous model parameters to minimize the sum of squares of the errors between predicted and experimental values. Then the second step was using ANNs to predict the remaining errors of the optimized mechanistic model. A hybrid model was the result of the coupling of the mechanistic model with the ANN model. The neural network models were the error predictors of the key variables of the mechanistic model. A three layer feed-forward neural network was developed using the sigmoid transfer function. The size of the hidden layer was varied between 3 and 6 neurons and the optimum value was retained. For the hybrid model, a neural network was developed for each of the five key variables describing the activated sludge process, namely the suspended solids, total chemical oxygen demand (COD_T) and ammonium (NH_4^+) of final effluent, dissolved oxygen (DO) in mixed liquor, and volatile suspended solids (VSS) in the returned activated sludge. For each variable, the inputs of the neural network models were selected among the available process variables using a cross-correlation algorithm between input and output variables, and a blend of trial and error and good judgment. Hourly data were used in training and testing the network. The resulting hybrid model was found, in general, to simulate with an adequate accuracy the dynamics of the activated sludge process.

Hack and Kohne (1996) used a feed-forward neural network to correlate process parameters of interest with simple and directly measurable parameters. They tried to

compensate for the breakdowns that occur in a process analyzer, that is part of a control loop, by estimating the respective process parameter using values of other parameters that were believed to have strong correlation with the process parameter of interest. In order to estimate the plant influent chemical oxygen demand (COD) and ammonium ($\text{NH}_4\text{-N}$) concentrations, the strength of the correlation, that exists between each of these two parameters and auxiliary parameters, such as conductivity, turbidity, nitrate, pH, redox potential, flow rate, and water temperature, were examined. A database that consisted of hourly sampling for a period of two consecutive weeks was used. One week was used for training the network and the other was used for verification. The back-propagation algorithm was used in training the network. By comparing the neural network estimates for influent COD and $\text{NH}_4\text{-N}$ with measured data for the testing period, the relative deviations were found to be 7.4% and 9.2% respectively.

Zhao et al. (1997) investigated the application of a hybrid model that integrates a mechanistic model and ANN to biological phosphorus removal in a sequencing batch reactor (SBR). A bench-scale SBR was used for generating the data and samples were taken at 15 minutes intervals. A simplified version of the Activated Sludge Model No. 2 (ASM2) was used. As a result of the reduced number of parameters in the simplified model, the calibration task was much easier, however, when testing the model with another data set, the predictions were in error. A neural network was integrated with the simplified model in order to improve the prediction capabilities of the hybrid model by learning how to bias the mechanistic model towards the measured values. A feed-forward neural network with 6 nodes in the hidden layer was utilized. Among the different activation functions, the sigmoid function was chosen for the hidden neurons. After training the hybrid model, the neural network error estimator improved the predictions.

1.2.3 The Box-Jenkins Transfer Function Methodology vs. Artificial Neural Networks (ANNs)

The main advantage of using the ANN modelling technique over the Box-Jenkins methodology is that ANNs have the ability to model nonlinearities in a system. However,

in order to obtain adequate predictions by an ANN model, a relatively large set of historical data is needed for training. The Box-Jenkins transfer function methodology offers the following advantages over ANNs: (1) a systematic approach for model identification using statistical measures such as cross-correlations between system inputs and outputs; (2) the statistical concepts that form the basis of the Box-Jenkins methodology allow the estimation of the confidence limits for the model predictions; and (3) the methodology offers a comprehensive system of model verification by the means of parameter and residual diagnostic checks. One disadvantage of the Box-Jenkins transfer function methodology (and any time series analysis method for that matter), is that the data set used has to have no gaps.

1.3 STUDY AREA – THE GOLD BAR WASTEWATER TREATMENT PLANT (GBWWTP)

The Gold Bar Wastewater Treatment Plant (GBWWTP) is the largest plant in the Edmonton area. The plant was constructed in 1956 on the southwest shore of the North Saskatchewan River. The present capacity of the plant is 950 ML/d for primary treatment and 420 ML/d for secondary treatment based on both hydraulic and process capacities. Approximately 95% of the sewage from the City of Edmonton are treated at the GBWWTP. The plant is typical of many conventional activated sludge plants designed for carbonaceous BOD and suspended solids reduction. It provides both primary and secondary treatment for the incoming raw sewage. Primary treatment consists of grit removal, mechanical screening, and primary sedimentation. The secondary treatment provides biological treatment in a suspended growth activated sludge system, final settling and microorganism reduction.

1.4 PROBLEM DESCRIPTION

The main scope of this research project was to develop dynamic models that can describe the following: (1) the wastewater inflow entering the plant; (2) the dynamic nature of the primary sedimentation process at the plant; and (3) the activated sludge

process at the plant. The objective of the modelling effort was to develop models which have the potential of being used as part of a RTC system that can be designed to improve the performance of the different processes at the plant. These models had to be adaptive and predictive.

This research program was designed to achieve several objectives. Those objectives included:

- (1) development of an artificial neural network model that utilizes rainfall data collected from the different rain gauges in the Edmonton area in order to provide short-term predictions of the wastewater inflow volume entering the plant;
- (2) studying the dynamics of the primary sedimentation process at the plant, and developing models that can describe the process and are able to provide short-term predictions of the quality parameters of the effluent wastewater leaving the process; and
- (3) studying the dynamics of the activated sludge process at the plant and developing a hybrid model that consists of two components: (a) a mechanistic component that is a simplified version of the Activated Sludge Model No. 1 (ASM1) developed by the International Water Association (IWA; formerly the International Association on Water Quality IAWQ); and (b) a black-box error-predictor component that can predict the error between the predictions by the mechanistic component and the actual experimental data.

1.5 ORGANIZATION OF THIS DISSERTATION

This dissertation is paper-format. In Chapter 2, the development of a wastewater inflow model utilizing ANNs has been presented. The development, training, and validation of the neural network model have been discussed in detail.

In Chapter 3, the development of neural network models for a full-scale primary sedimentation tank at the GBWWTP has been presented. Description of the sampling

program, development, training, and validation of the models have been discussed in detail.

In Chapter 4, the development of transfer function models for the primary sedimentation tank has been presented. Model identification, estimation, and validation have been discussed in detail.

In Chapter 5, the development of a simplified version of the ASM1 for a full-scale activated sludge process at the GBWWTP has been presented. Description of the sampling program, model development, and model testing have been discussed in detail.

In Chapter 6, the development of a hybrid model, that is composed of a deterministic and ANN/time series components, and describes the activated sludge process has been presented. Model development, training, and validation have been discussed in detail.

In Chapter 7, the conclusions on the modelling effort and model performance has been presented. Finally, suggestions on future research along these topics have been recommended.

1.6 REFERENCES

1. Barnes, W.J. and F.A. Rowe, "Modelling Sewer Flows Using Time Series Analysis", J. Env. Eng., ASCE 104(EF4): 639-646 (1978).
2. Beck, M.B., "System Identification and Control", In *Dynamic Modelling and Expert Systems in Wastewater Engineering* (MI, USA: Lewis Publishers, Inc., 1989).
3. Berthouex, P.M., W.G. Hunter, L. Pallesen, and C.Y. Shih, "Modelling Sewage Treatment Plant Input BOD Data", J. Env. Eng., ASCE 104(EF1): 127-138 (1975).
4. Box, G.E. and G.M. Jenkins, *Time Series Analysis: Forecasting and Control* (Oakland, CA, USA: Holden-Day, 1976).
5. Bruce, R.A. and G.G. Party, "Knowledge-Based System for the Diagnosis of an Activated Sludge Plant", In *Dynamic Modelling and Expert Systems in Wastewater Engineering* (MI, USA: Lewis Publishers, Inc., 1989).
6. Capodaglio, A.G., H.V. Jones, V. Novotny, and X. Feng, "Sludge Bulking Analysis and Forecasting: Application of System Identification and Practical Neural Computing Technologies", Wat. Res. 25: 1217-1224 (1991).
7. Capodaglio, A.G., S. Zheng, V. Novotny, and X. Feng, "Stochastic System Identification of Sewer-Flow Models", J. Env. Eng., ASCE 116(EF2): 284-298 (1990).
8. Caroni, E., V.P. Singh, L. Ubertini, "Rainfall Runoff Sediment Yield Relationship by Stochastic Modelling", Hydrological Science Journal 29(2): 203-218 (1984).
9. Chow, K.C.A., W.E. Watt, D.G. Watt, "A Stochastic-Dynamic Model for Real Time Flood Forecasting", Water Resources Research 19(3): 746-752 (1983).
10. Cote', M., P.A. Grandjean, P. Lessard, and J. Thibault, "Dynamic Modelling of the Activated Sludge Process: Improving Prediction Using Neural Networks", Wat. Res. 29:995-1004 (1995).
11. Du, Y.G., T.R. Sreekrishnan, R.D. Tyagi, and G.C. Campbell, "Effect of pH on Metal Solubilization from Sewage Sludge: A Neural-Net-Based Approach", Can. J. Civil Eng. 21: 728-735 (1994).

12. Gall, R.A.B. and G.G. Patry, "Knowledge-Based System for the Diagnosis of an Activated Sludge Plant", In *Dynamic Modelling and Expert Systems in Wastewater Engineering* (MI, USA: Lewis Publishers, Inc., 1989).
13. Gurnell, A.M. and C.R. Fenn, "Box-Jenkins Transfer Function Models Applied to Suspended Sediment Concentration-Discharge Relationship in a Proglacial Stream", *Artic and Alpine Research* 16(1): 93-106 (1984).
14. Hack, M. and M. Kohne, "Estimation of Wastewater Process Parameters Using Neural Networks", *Wat. Sci. Tech.* 33(1): 101-115 (1996).
15. Hoff, J.C. *A Practical Guide to Box-Jenkins Forecasting* (CA, USA: Lifetime Learning Publications, Inc. 1983).
16. Huck, P.M. and G.J. Farquhar, "Water Quality Models Using the Box-Jenkins Method", *J. Env. Eng., ASCE* 100(EF3): 733-752 (1974).
17. Jayawardena, A.W. and F. Lai, "Time Series Analysis of Water Quality Data in Pearl River, China", *J. Env. Eng., ASCE* 115(EF3): 590-607 (1989).
18. Katebi, R., M.A. Johnson, and J. Wilkie, *Control and Instrumentation for Wastewater Treatment Plants* (London, UK: Springer-Verlag London Limited, 1999).
19. Lee, D.M., *Application of ARIMA Models to Wastewater Treatment Data: M.Sc. Thesis* (Milwaukee, Wisconsin, USA: Marquette University, 1986).
20. Lemke, K.A., "Transfer Function Models of Suspended Sediment Concentration", *Water Resources Research* 27(3): 293-305 (1991).
21. Lessard, P. and M.B. Beck, "Dynamic Modelling of the Activated Sludge Process: A Case Study", *Wat. Res.* 27: 963-978 (1993).
22. Novotny, V., A. Capodaglio, and H. Jones, "Real Time Control of Wastewater Treatment Operations", *Wat. Sci. Tech.* 25(4-5): 89-101 (1992).
23. Olsson, G., "Practical Experiences of Identification and Modelling from Experiments", In *Dynamic Modelling and Expert Systems in Wastewater Engineering* (MI, USA: Lewis Publishers, Inc., 1989).
24. Premier, G.C., R. Dinsdale, A.J. Guwy, F.R. Hawkes, D.L. Hawkes, and S.J. Wilcox, "A Comparison of the Ability of Black Box and Neural Network Models of

- ARX Structure to Represent a Fluidized Bed Anaerobic Digestion Process", *Wat. Res.* 33(4): 1027-1037 (1999).
25. Pu, H.C. and Y.T. Hung, "Prediction of a Medium-Sized Municipal Wastewater Treatment Plant Performance with Artificial Neural Networks", *Water Treatment* 9: 305-322 (1994).
 26. Rossman, L.A., "Applications of Expert Systems in Environmental Engineering", In *Dynamic Modelling and Expert Systems in Wastewater Engineering* (MI, USA: Lewis Publishers, Inc., 1989).
 27. Rumelhart, D.E., G.E. Hinton, and J.L. McClelland, *A General Framework for Parallel Distributed Processing. Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (Cambridge, MA, USA: MIT Press, 1: 45-76, 1986).
 28. Su, H. T. and T.J. McAvoy, "Long-Term Prediction of Chemical Processes Using Recurrent Neural Networks: A Parallel Training Approach", *Ind. Eng. Chem. Res.* 31: 1338-1352 (1992).
 29. Tabak, H.H. and R. Govind, "Prediction of Biodegradation Kinetics Using a Nonlinear Group Contribution Method", *Environmental Toxicology and Chemistry* 12: 251-260 (1993).
 30. Thompstone, R.M., K.W. Hipel, A.I. McLeod, "Forecasting Quarter Monthly Riverflow", *Water Resources bulletin* 21(5): 731-741 (1985).
 31. Tyagi, R.D. and Y.G. Du, "Kinetic Model for the Effects of Heavy Metals on Activated Sludge Process Using Neural Networks", *Env. Tech.* 13: 883-890 (1992).
 32. Tyagi, R.D., Y.G. Du, T.R. Sreekrishnan, and J.P. Villeneuve, "Neural network model for the operational control of activated sludge processes", *Process Biochemistry* 28: 259-267 (1993).
 33. Wilcox, S.J., D.L. Hawkes, F.R. Hawkes, and A.J. Guwy, "A Neural Network, Based on Bicarbonate Monitoring, to Control Anaerobic Digestion", *Wat. Res.* 29: 1465-1470 (1995).
 34. Zhao, H., O.J. Hao, T.J. McAvoy, and C.H. Chang, "Modelling Nutrient Dynamics in Sequencing Batch Reactor", *J. Env. Eng., ASCE* 123(4): 311-319 (1997).

CHAPTER 2. WASTEWATER INFLOW PREDICTION USING AN ARTIFICIAL NEURAL NETWORK*

2.1 INTRODUCTION

The wastewater inflow rate entering a treatment facility depends on the characteristics of the area and population served by the sewer system discharging to the plant and the occurrence of meteorological events. The influent wastewater flow may increase substantially during rainfall events. Variations in flow of over two orders of magnitude are not uncommon in combined sewer systems (Capodaglio, 1994). As a result, during such events, the hydraulic capacity of the plant may be exceeded and upsets to the biological processes may occur. Short-term predictions of the wastewater inflow rate during rainfall events would be essential if real-time control of different unit processes in a treatment plant is to be utilized.

Hydraulic load prediction is usually derived from hydrological (rainfall/runoff) and hydraulic (sewer flow routing) simulation models. The majority of these models that are developed to date are deterministic. Such models require detailed knowledge of the drainage area under study and usually incorporate a large number of parameters that may vary in time following a stochastic dynamic pattern, and therefore, calibrating this type of model is a very tedious job. Although deterministic models provide a solid foundation for the understanding of the physical phenomena, their characteristics make them inadequate for application in real-time control (Capodaglio, 1994). Another type, or group, of models is the so-called "black-box" class of models. These models do not try to describe the internal functions of a system; however, they can recognize relationships that exist between measurable inputs and outputs of the system. In a recent study conducted by Carstensen et al. (1998), prediction of the hydraulic load to a treatment plant one hour ahead was done using three different models. The models represent three different levels of complexity, ranging from a simple regression model over an adaptive grey-box model (composed of a simple deterministic regression model and a stochastic model

A part of this chapter has been published. Gamal El-Din, A. and D.W. Smith, *Water Research* 36: 1115-1126 (2002).

Another part of this chapter has been accepted for publication. Gamal El-Din, A. and D.W. Smith, *Journal of Environmental Engineering and Science* (April 2002).

component) to a complex hydrological model, which was calibrated based on extensive measurement campaigns in the sewer system. They showed that the simple regression model and the grey-box model performed better than the complex hydrological flow model and they recommended that for model-based predictive control, models should be kept simple and identifiable from measured data. Among the "black-box" class of models are artificial neural network (ANN) models. Presented in this paper, a case study for the Gold Bar Wastewater Treatment Plant (GBWWTP) in Edmonton, Alberta, Canada, in which an ANN model was built and used to provide short-term predictions of wastewater inflow rate at the entrance to the plant.

2.2 DESCRIPTION OF THE AREA UNDER STUDY

In 1956, the GBWWTP was constructed on the south bank of the North Saskatchewan River. The plant is typical of many conventional step-feed activated sludge plants designed for carbonaceous BOD and suspended solids removal. It provides grit removal, screening, primary settling, biological treatment in a suspended growth activated sludge system, final settling, and microorganism reduction. The initial dry-weather flow capacity of the plant was 91 ML/d, serving the needs of a population of 250,000. The present capacity of the plant is 950 ML/d for primary treatment and 420 ML/d for secondary treatment based on both hydraulic and process capacities. The plant treats domestic and industrial sewage from the City of Edmonton. Approximately 95% of the sewage flows from the City are treated at the GBWWTP.

The City of Edmonton's sewerage and drainage facilities include 1619 km of sanitary sewer, 934 km of combined sewer, and 1808 km of storm sewer. The City's sewerage system covers a drainage area of 35,768 ha and serves a total population of 610,000. Figure 2.1 shows a schematic of the drainage area and the location of the plant. The City of Edmonton's Drainage Branch operates a network of approximately 30 permanent flow monitors in the sewerage system. The GBWWTP is one of the permanent flow/water-level installations, which are supplied with power and are equipped with a modem so that data can be downloaded from a remote location. The GBWWTP uses a

measuring flume in conjunction with a water level sensor. The City of Edmonton operates a network of 18 tipping bucket rain gauges. These gauges are shown in Figure 2.1 and are numbered according to their site identification numbers in the maps provided from the City of Edmonton Drainage Branch. These rain gauges are normally in operation for the period from May through September. The incidence of a tip of the bucket is recorded by the rain gauge, which represents 0.2 mm of rainfall since the last bucket tip occurred. Rain gauges are calibrated prior to re-installation each May. Each rain gauge is connected to a data logger on site, and each site is downloaded automatically by modem from Branch headquarters.

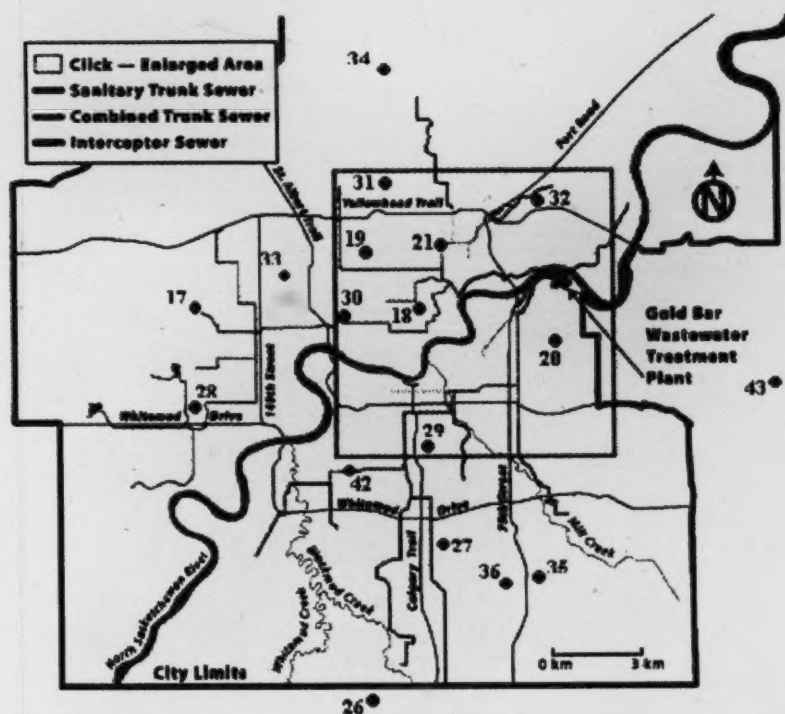


Figure 2.1 The City of Edmonton drainage area.

2.3 OVERVIEW OF ANN

Artificial neural networks go by many names such as connectionist models and parallel distributed processing models. The ANN modelling approach is a computer

methodology that has the ability to apply information gained from past experience to new problems or case scenarios. An ANN uses many simple computational elements, named artificial neurons, connected by variable weights. Although each neuron, alone, can only perform simple computations, the hierarchical organization of a network of interconnected neurons makes an ANN capable of performing complex tasks such as pattern classification and prediction. Among the many different architectures, the multi-layer perceptron architecture is commonly used for prediction. In multi-layer perceptron architecture, neurons are arranged in layers, an input layer, hidden layers, and an output layer. Figure 2.2 shows a schematic diagram of a multi-layer neural network.

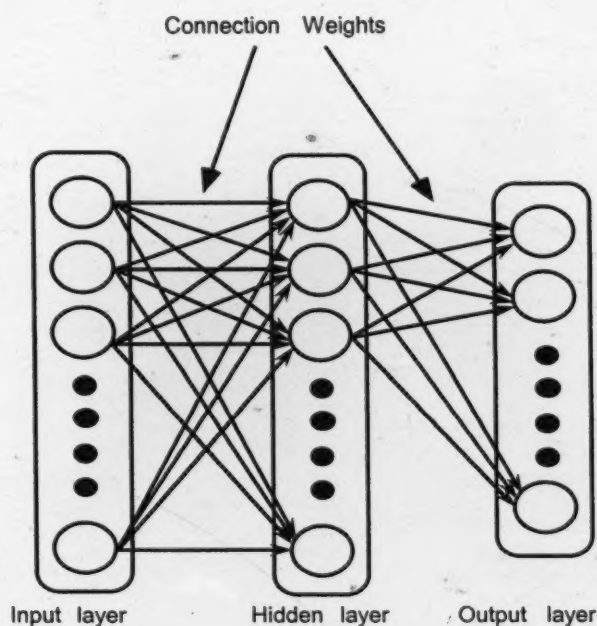


Figure 2.2 A schematic diagram of a multi-layer neural network.

Input neurons, or nodes, receive values of an instance of the input parameters that are fed to the network after being scaled into a numeric range that is efficient for calculations by the neural network model. Outcomes to the output parameters for the instance under consideration are assigned by the output neurons. Hidden neurons connect the input neurons to the output neurons and provide nonlinearity to the network. Each neuron is connected to every neuron in adjacent layers by a connection weight, which

determines the strength of the relationship between two connected neurons. The output from a neuron is multiplied by the connection weight before being introduced as input to the neuron in the next layer. Each neuron, except those in the input layer, sums all of the inputs that it receives and the sum is converted to an output value according to a predefined activation, or transfer, function.

For prediction problems, a supervised learning algorithm is often utilized for teaching the network how to relate input data patterns to output data. The main reason for the popularity of the multi-layer perceptron architecture in recent years is the development of new training algorithms like the back-propagation algorithm (Lippmann, 1987). Since its development, the back-propagation algorithm has been widely used for training multi-layer neural networks. The algorithm uses a gradient search technique to minimize a cost function equal to the mean square difference between the desired and the actual net outputs. It requires a continuous differentiable nonlinearity to be used as the transfer function by the neurons. The algorithm consists of the following steps: (1) all weights are assigned initial small random values; (2) input data are propagated in a feed-forward manner through the network to produce output data according to the weights and the transfer function; (3) the outputs produced are compared with the target outputs which are known in advance; (4) the errors generated are then propagated backwards in a certain manner through the network for adjustments of the present weights using two factors, namely, a learning factor and a momentum factor; and (5) this procedure is repeated for each training example in the training set; a cycle (an epoch) represents one pass over the whole training set; multiple epochs are required until a satisfactory data mapping is achieved. A mathematical description of the training algorithm is provided elsewhere (Rumelhart et al., 1986 1, 2). The learning factor (ranges from 0 to 1) represents the step size by which the weights are updated and is problem-specific. For noisy data, it is better to keep it below 0.1 (Gallant, 1993). In order to speed convergence of the back-propagation algorithm, a momentum factor (also ranges from 0 to 1) can sometimes be used. The idea is to keep weight changes on a faster and more even path by adding fractions of previous weight changes. A reasonable value for the momentum factor is 0.9 (Gallant, 1993).

2.4 DEVELOPMENT OF THE MODEL

2.4.1 Input and Output Variables

The neural network model was built to provide in-advance prediction of the quantity of wastewater flow that enters the treatment plant during rainfall events. The selection of the input variables had to be done in a way that enables the neural network model to accomplish this task. Among the 18 rain gauges that cover the major drainage basins of the City's sewerage system, 8 rain gauges (gauges #18, 19, 20, 21, 29, 30, 31, and 32) were chosen to be used by the neural network. As shown in Figure 2.1, these gauges were selected because they are located in drainage areas that have high percentage of combined sewers. An index to represent the day-of-the-week and another index to represent the hour-of-the-day were used as inputs to the neural network model, and this way, the model can learn the diurnal pattern of wastewater flow entering the plant on different days of the week during dry weather flow conditions, and hence, the network can have the ability to differentiate between two storms having very close characteristics but occurring at different times.

2.4.2 Source Data Screening and Analysis

Records of wastewater raw influent entering the plant were downloaded from the City of Edmonton's Drainage Branch computer network for the summers (May-September) of 1995 to 1997. Flow records were at 10-minute intervals. The raw data flow records had some negative values, which indicated segments of faulty data. By visually examining these periods of faulty flow data, it was found that these periods contained sequences of very high, very low, and negative values of flow and it was easy to identify the starts and ends of these periods of faulty data. These segments of data were eliminated, and hence, were considered as data gaps. Flow records were averaged over 30-minute intervals. As it was mentioned before, the rain gauges used by the City of Edmonton's Drainage Branch are tipping bucket rain gauges that record the incidence of a tip of the bucket, which represents 0.2 mm of rainfall since the last bucket tip occurred. As a result, the data had to be arranged in a way that can be understood by the neural

network model. In order to do that, the accumulative rainfall recorded by each of the 8 rain gauges over 30 minute periods was calculated. A rainfall data record at time t refers to the amount of rain accumulated from $t-0.5$ hours to t hours, which means the amount of rain collected during the past 30 minutes.

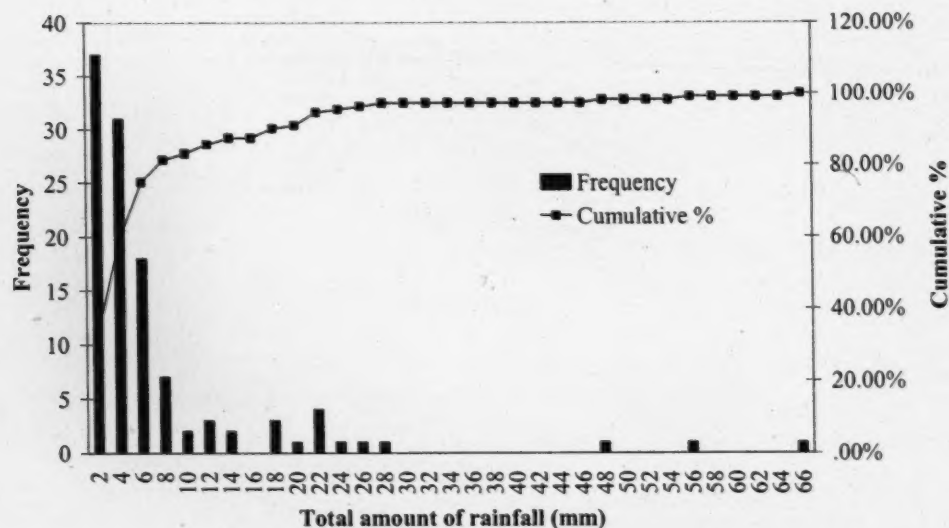


Figure 2.3 A histogram of the total amount of rainfall recorded during rainfall events.

The objective of the modelling effort was to predict the inflow to the treatment facility during rainfall events. The data set used in the present study contained 114 rainfall events. These events differed substantially in both the total amount of rainfall as well as the maximum intensity. Out of the 114 rainfall events that were encountered in the data set, 90 events caused the secondary capacity of the plant to be exceeded and secondary bypass had to be used during those events. Figure 2.3 shows a histogram of the total amount of rainfall (averaged for the 8 gauges) recorded for the rainfall events encountered in the data set. The total amount of rainfall ranged from 0.5 to 64.8 mm. Around 75% of the rainfall events had less than 6 mm of total rainfall; 16% had total rainfall in the range 6 to 20 mm. Only 10 rainfall events had total amount of rainfall that exceeded 20 mm; of these events 3 storms had total rainfall over 45 mm. Each of these 3

events lasted more than 15 hours with the one that had the maximum total amount of rainfall of 64.8 mm lasting about 46 hours.

After training a neural network on historical data, the network has to be tested against a data set (usually called a testing or validation or verification set) that was never been seen before by the network during the course of training. The ability of the network to generalize from the training data set to the validation data set will determine its performance. Therefore, a thorough source data analysis has to be conducted when dividing the available data set into the two parts for training and validation. The validation data set has to be within the range of the training set, otherwise, the network will extrapolate beyond the scope that it has been trained for. Extrapolation may give erroneous results, and hence, the performance of the neural network model will be misjudged.

In the present study, the important features in the data that the network should recognize are the rainfall events. The available data set contained 114 rainfall events and was divided into three sets, namely, set #1, set #2, and set #3 as shown in Figure 2.4. Data set #1 was always used as part of the training data set. Data sets #2 and #3 were used as validation data sets and each contained 22 rainfall events. It is clear from Figure 2.4 and Figure 2.5 that each of the two validation data sets included rainfall events that covered a wide range of total amount of rainfall, maximum intensity, maximum flow at the entrance to the plant that was encountered during the rainfall event, and total inflow (amount of flow, conveyed to the plant during a rainfall event, that was above the normal dry weather flow).

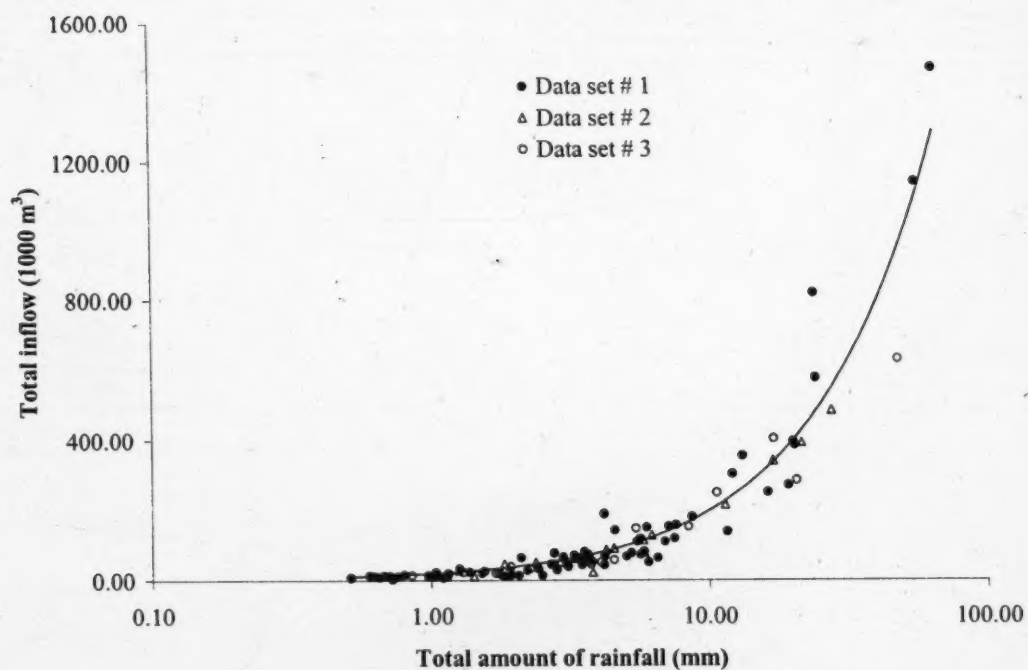


Figure 2.4 Total inflow above dry weather flow.

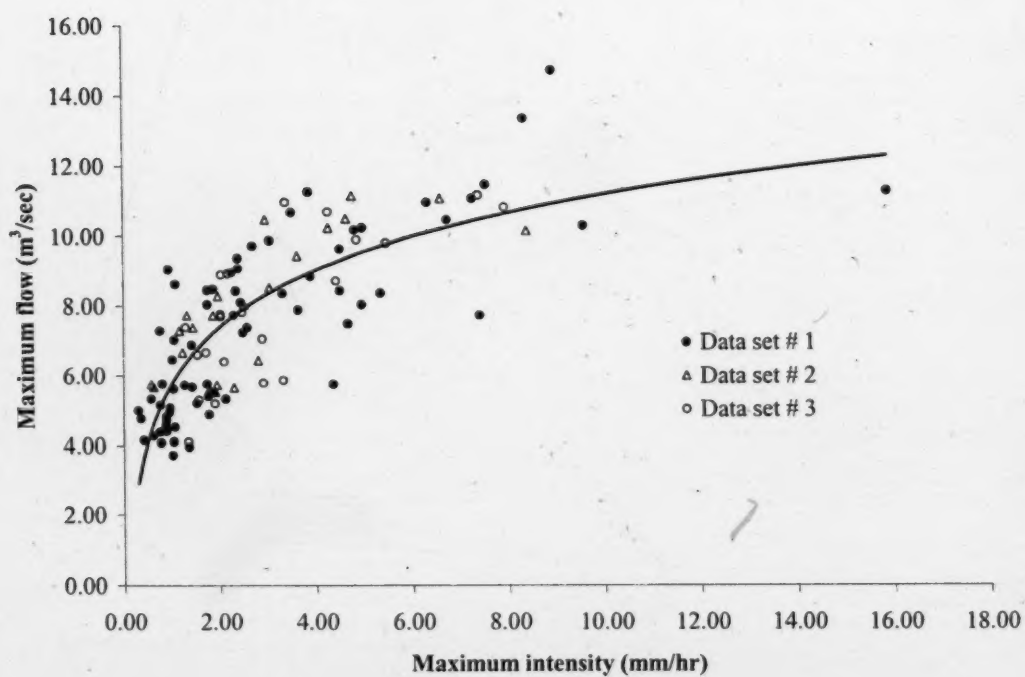


Figure 2.5 Maximum flow encountered at the entrance to the plant.

2.4.3 Architecture of the Neural Network Model

Much of the success in applying neural networks to engineering problems is mainly attributed to the achievements of the multi-layer feed-forward neural network architecture with back-propagation training algorithm (Lippmann, 1987; Tyagi and Du, 1992; Cote' et al., 1995; Zhao et al., 1997; Djebbar and Kadota, 1998). In this ANN architecture, information is processed in a forward manner through the network while the prediction error is propagated backwards through the network. A feed-forward neural network with back-propagation training algorithm was used for the modelling that was conducted in this study. As was mentioned before, the day-of-the-week index, the hour-of-the-day index, and rainfall data were the input parameters to the neural network model. Seven input nodes were used for day-of-the-week indicators, each of which corresponds to one weekday. If the day was Sunday, the corresponding node takes a value of "1" and all other 6 nodes take the value of "0". The hour-of-the-day index was modelled using 24 input nodes. Similarly to day-of-the-week indicators, only one hour-of-the-day indicator would take the value of "1" at any time step, with the other 23 nodes having a value of "0".

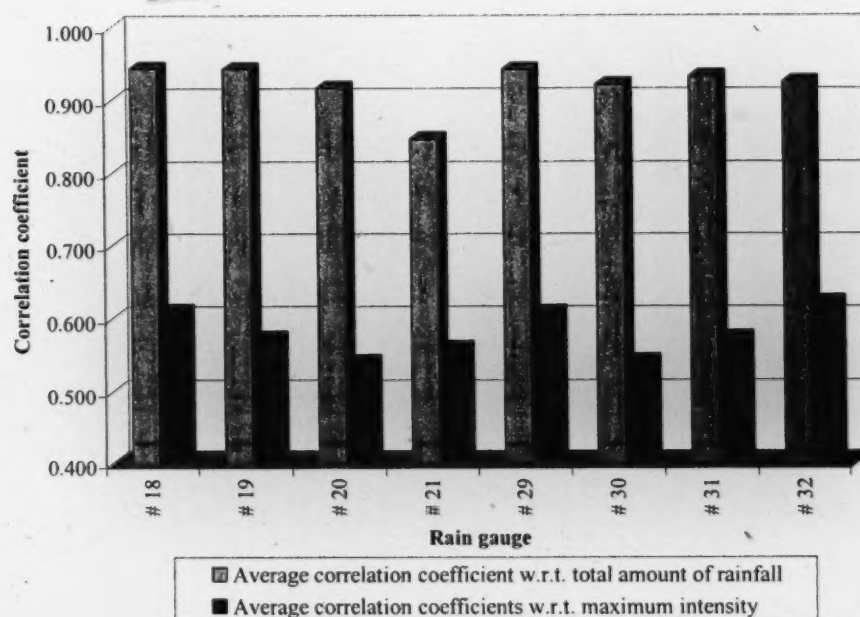


Figure 2.6 Maximum flow encountered at the entrance to the plant.

After analyzing the rainfall data recorded by the 8 gauges during the 114 rainfall events encountered in the present study, a correlation matrix was calculated for the 8 gauges with respect to the maximum intensity recorded during the event as well as the total amount of rainfall recorded. For each of the 8 gauges, the 8 correlation coefficients were averaged and the results are shown in Figure 2.6, from which it is clear that the variability among the 8 gauges with respect to the maximum intensity recorded during a rainfall event is almost two times the variability with respect to the total amount of rainfall recorded. However, this is expected for such a large drainage area.

From Figure 2.6 and from the visual examination of the rainfall data, recorded by each of the 8 rain gauges, it was clear that most of the 8 gauges recorded rainfall during most of the events that occurred. However, the intensity of rainfall events differed substantially from one gauge to another. During some rainfall events, the intensity of the rainfall recorded by different gauges varied by almost two orders of magnitude. During a few of the rainfall events, some of which lasted for more than 15 hours, a few gauges did not record any rainfall. This simply could be due to a malfunction of the gauge during the event or due to the spatial variations in the rainfall on such a large drainage area. Because of the non-homogenous characteristics of rainfall events encountered, it would be misleading to use only one gauge to represent the whole drainage area serviced by the treatment plant.

For the sake of simplifying the neural network model, the rainfall data records that were recorded by the 8 gauges were averaged and the averaged data were used to reduce the dimensionality of the input layer of the neural network. It is evident from Figure 2.6 that gauge #21 showed less agreement with the other 7 gauges. From the source data analysis, it was found that the reason for that was a malfunction of the gauge for almost one and half months of the summer of 1997. During this period, 18 rainfall events had occurred that were not recorded by gauge #21, but were recorded by all the other 7 gauges. Despite this fact, the data for gauge #21 were utilized without any modifications in order to demonstrate that using averaged rainfall data of the 8 gauges has the advantage of reducing the effect of a malfunction, that may occur to one or more gauges

in a real-life situation, on the predictions of a neural network model that can be utilized online as a part of a real-time control system.

In order to present the rainfall data records to the neural network, a moving window of past records had to be used, however, the size of the moving window was still to be determined. Using a very small window may hinder the ability of the network to provide adequate predictions because the whole picture is not fully seen by the neural network model. On the other hand, if a very large window size was to be used, the dimensionality of the input layer will increase which will necessitate the utilization of more parameters. In neural networks, this means more hidden nodes, and as a result, more connection weights. In the present study the size of the moving window was changed incrementally from 3 hours horizon (6 rainfall records) to 10 hours (20 rainfall records) with increments size of one hour. This, as will be shown later, allowed the selection of the time span that gives the best predictions and, at the same time, kept the model as simple as possible.

Using these input parameters, predictions of the wastewater flow entering the plant 30-minute ahead were made by the neural network, and hence, the output layer consisted of only one node. If the neural network model was to be used as a part of a real-time control system, this forecast horizon would allow an operator on duty to take an action like bringing a primary sedimentation unit back to service or changing an activated sludge unit from plug flow mode to step-feed mode, which are two normal procedures that are used by the operators at the Gold Bar Wastewater Treatment Plant in periods of wet weather flows. However, expanding the forecast horizon can be easily conducted by retraining the neural network model to adapt with the new forecast horizon. For example, expanding the forecast horizon from 30-minute to one hour will shift the moving window one rainfall record.

In large part, the architecture of a neural network is determined by the format selected for its inputs and outputs, however, still to be determined is how many hidden layers should be used in the network and of what sizes should they be. Many researchers in the field of artificial neural networks suggest that it is usually unnecessary to use more

than one hidden layer in a multi-layer feed-forward network and varying the number of hidden nodes in the one hidden layer is usually sufficient for delivering distinct results. Most applications of back-propagation to neural network problems make use of only one hidden layer. A major reason for this is that intermediate cells not directly connected to output cells will have very small weight changes and will learn very slowly (Gallant, 1993). In the present study only one hidden layer was used.

Although the relationship between network performance and the size of hidden layer is not well understood, there are some principles which can be used as a guide. Among these, the most important one is the principle of generalization versus convergence, two aspects of network behavior, which often work against one another. Generalization means the ability of the network to produce reasonable results when applied to a data set that has not been seen before by the network during the course of training. Convergence is simply the ability of the network to learn the training data. In general, the likelihood of convergence will increase when more hidden neurons are used, however, using too many hidden nodes will make the network generalize poorly, memorizing the training data rather than focusing on its significant features. This principle is similar to the principle of fitting versus forecasting in time-series analysis. The objective is to use as many hidden nodes as needed for convergence without inhibiting the ability of the network to generalize, so the network would be able to focus on the important features in the data rather than fitting the noise that is an inherent component of any environmental field data set. In the present study, from the initial testing of the model, it was found that the smallest size of the hidden layer that allowed the network to converge was 3 nodes. The size of the hidden layer was increased incrementally from 3 nodes to 10 nodes with an increment size of one node and the parsimonious model that gave the best results was selected.

2.4.4 Training the Network

In the modelling that was undertaken in this study, a batch-mode (weight updates were done after each epoch and not after each training pattern) back-propagation

algorithm was used in the course of training and the network was saved at the point of minimum training error. This mode of training is insensitive to the selection of the value of the learning and momentum factors, and hence, was used in the present study to reduce the dimensionality of the variables' space when the network is to be trained. Training was conducted using NeuroShell 2 software from Ward Systems Group Inc.

Before data patterns can be presented to a neural network, they have to be scaled into a range that is suitable for the calculations by the training algorithm. In the present study, data were scaled linearly into the range $<<0,1>>$. Because of its favorable characteristics, the logistic function is perhaps the most widely used transfer function with the back-propagation algorithm. From the initial model testing conducted in the present study using three different transfer functions by the hidden layer (logistic, tanh, and Gaussian), the logistic function was found to yield the best results. A linear, rather than a logistic, function in the output node was used as the initial testing of the model showed better results when using a linear function in the output node.

A big dilemma, which is always involved in training feed-forward back-propagation neural networks, is when to stop training. In other words, after how many cycles (epochs) should training be stopped in order to allow the neural network to learn the important features from the data without memorizing them, so that it can generalize well when applied to the validation data set. In the present study training was stopped after 10, 20, 30, 40, 50, 75, 100, 150, 200, 300, 400, 500, 750, and 1000 epochs, for every candidate model that was tested. Each time training was stopped, the model was tested against the verification data set. It will be seen later that using this systematic approach, the point at which training should be stopped so that generalization could be maximized was identified.

The historical data set contained a total of 16,001 data records, which were a blend of dry and wet weather flow data records for the summers of 1995 to 1997. Around 80% of the data were dry weather flow records that were used only to enable the network to learn the dry weather flow patterns during different days of the week. The other 20% are

the wet weather flow patterns, which represent the important features in the data. The objective was to predict the flow entering the plant during wet weather flow conditions, and as a result, the performance of the candidate models (with respect to training and validation) was judged only using the wet weather flow patterns. As was mentioned before, the historical data set contained 114 rainfall events, which were divided into three sets, namely data sets #1, #2, and #3. When any of data sets #2 or #3 was used as the validation set the other two sets were used as parts of the training set.

2.5 RESULTS AND DISCUSSION

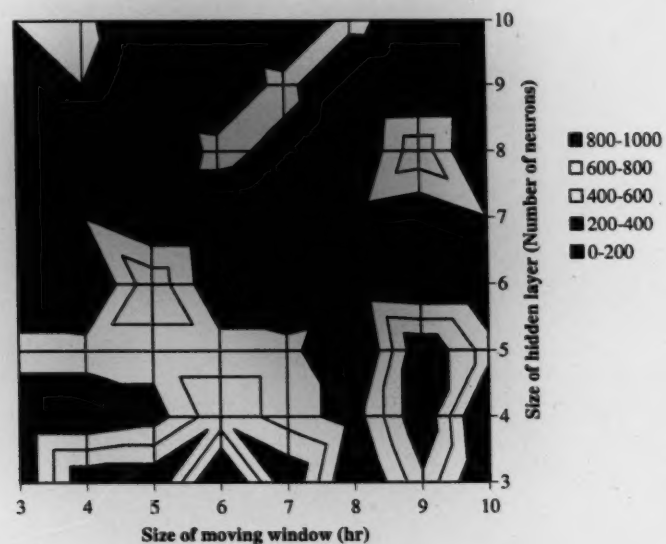
The performance of the network was measured using the coefficient of determination (R^2), which is a statistical indicator that compares the accuracy of the model to the accuracy of a trivial benchmark model wherein the prediction is just the mean of all the samples. The coefficient of determination, R^2 , is mathematically described as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{i=n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{i=n} (y_i - \bar{y})^2} \quad [2.1]$$

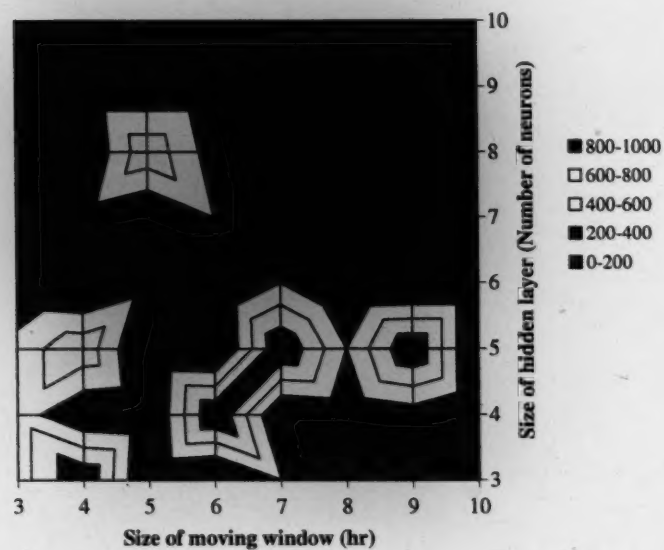
where y_i is the actual output value, $\hat{y}_i$ is the output value predicted by the network, $\bar{y}$ is the mean of y values, and n is the total number of data records. A perfect fit would result in an R^2 value of "1" while R^2 value of "0" means a very poor fit.

In order to try 8 different sizes of the moving window, 8 different sizes of the hidden layer and stopping training at 14 different number of epochs, a total of 896 ($8 \times 8 \times 14$) runs were conducted twice. Once when data set #2 was used as the validation data set (while sets #1 and #3 were used as part of the training data set) and the second when data set #3 was used for validation (while sets #1 and #2 were used as part of the training data set). For each size of the moving window, 8 different sizes of the hidden

layer were tried, and for each one of them, training was allowed for 14 different number of cycles.



(a) Data set #2 was used for validation



(b) Data set #3 was used for validation

Figure 2.7 The number of epochs at which the maximum value of the R^2 was obtained.

It was found that for almost all the candidate models that were able to converge, as the number of epochs that was used in training increased, the value of the R^2 for the training set increased and the maximum value was obtained when training was allowed for 1000 epochs (the maximum number of epochs that was used). On the other hand, the maximum value of the R^2 for the validation data set was always obtained at a much lower number of epochs (in one instance was as low as 30 epochs).

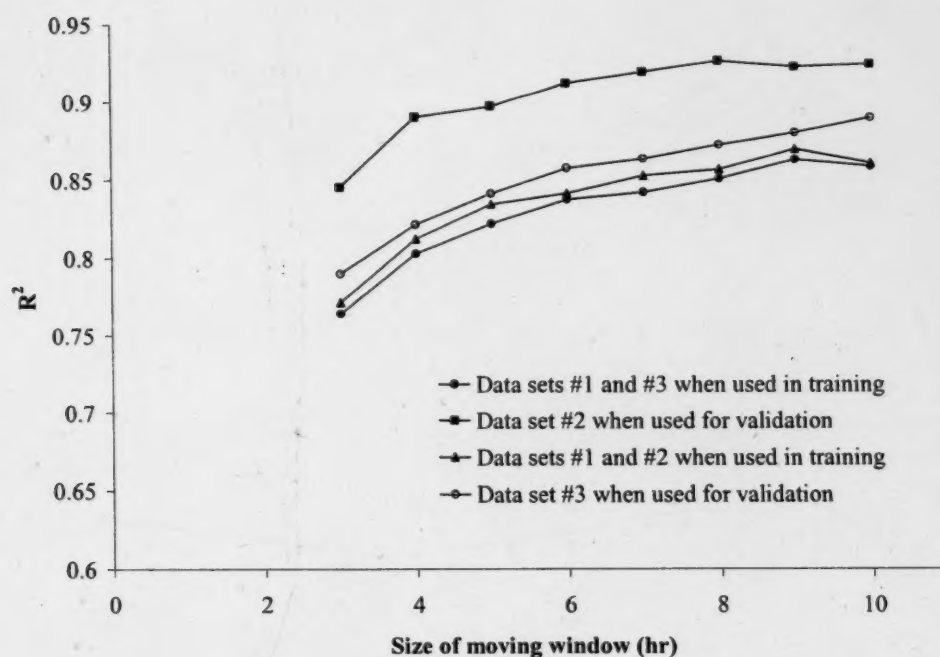


Figure 2.8 The effect of increasing the size of moving window.

Figure 2.7 shows the number of epochs, at which the maximum value of the R^2 for the validation data set was obtained. It is clear from Figure 2.7 that when the dimensionality of the network increased (either the size of the moving window or the size of the hidden layer is in the higher range) the maximum value of the R^2 for the validation data set was obtained at low number of epochs and allowing training beyond this point hindered the ability of the network to generalize.

Only when using a moving window of sizes 3 and 10 hours, were all the 8 candidate models able to converge to an adequate solution. However, for all the other sizes, one or more model could not converge to an adequate solution. For example, using a moving window of size 4, the performance of the network when using a hidden layer of size 3, 5, and 7 was much lower than that when using other sizes of the hidden layer. This could be because the neural network was trapped in a local minimum. Although increasing the dimensionality of a network increases its chances to converge, a network that is trained with a gradient descent algorithm, like the back-propagation, is not guaranteed to find a global minimum, even if it converges (Gallant, 1993).

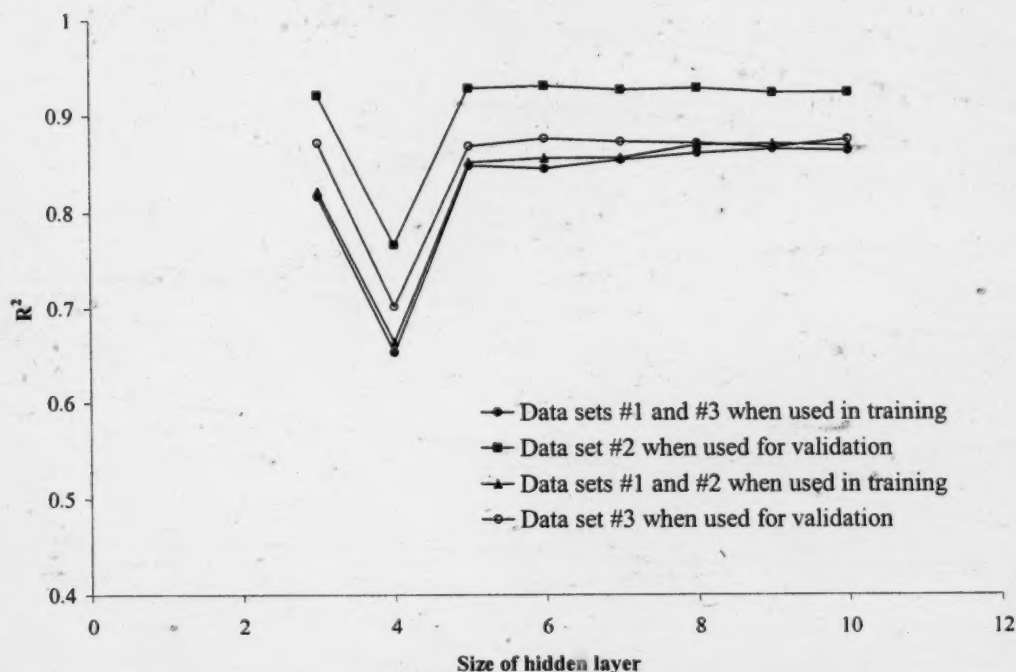


Figure 2.9 Results obtained using a moving window of 8-hr.

When averaging the results of the candidate models that were able to converge to an adequate solution, Figure 2.8 was obtained, which clearly shows that increasing the size of the moving window up to 8 hours had the effect of improving the performance of the network with respect to both training and validation. Expanding the moving window beyond 8 hours had very little effect on the performance. Therefore, and because a

parsimonious model is always the objective of any modelling effort, a moving window of size 8 hours was chosen. Figure 2.9 shows the results obtained using the 8 hours moving window with different sizes of the hidden layer. It is clear from Figure 2.9 that even though increasing the size of the hidden layer always improved the R^2 value for training, the generalization ability of the model, represented by the R^2 value of the validation set, did not have any substantial improvement beyond a hidden layer of size 5. It is also clear from Figure 2.9 that increasing the size of the hidden layer from 3 to 5 had a substantial improvement in the ability of the network to learn the training data, however, increasing the size beyond 5 nodes had a much lower effect. A network that utilizes an 8 hours moving window and a 5 node hidden layer was chosen to be the final model.

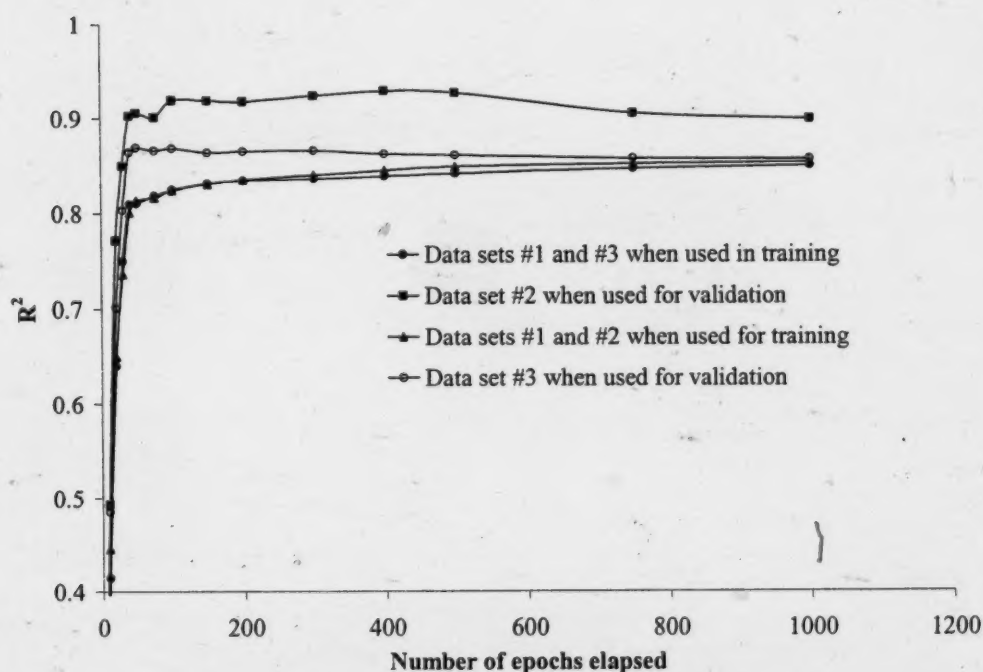


Figure 2.10 Results obtained by the final model.

Results obtained by the final model are shown in Figure 2.10, from which it is evident that continuing training beyond 400 epochs did not improve the value of the R^2 for validation and had a slight improvement in the value for the training set. Therefore, it

was decided to stop training of the final model after 400 epochs had elapsed. Dry weather flow patterns simulated by the final model are shown in Figure 2.11.

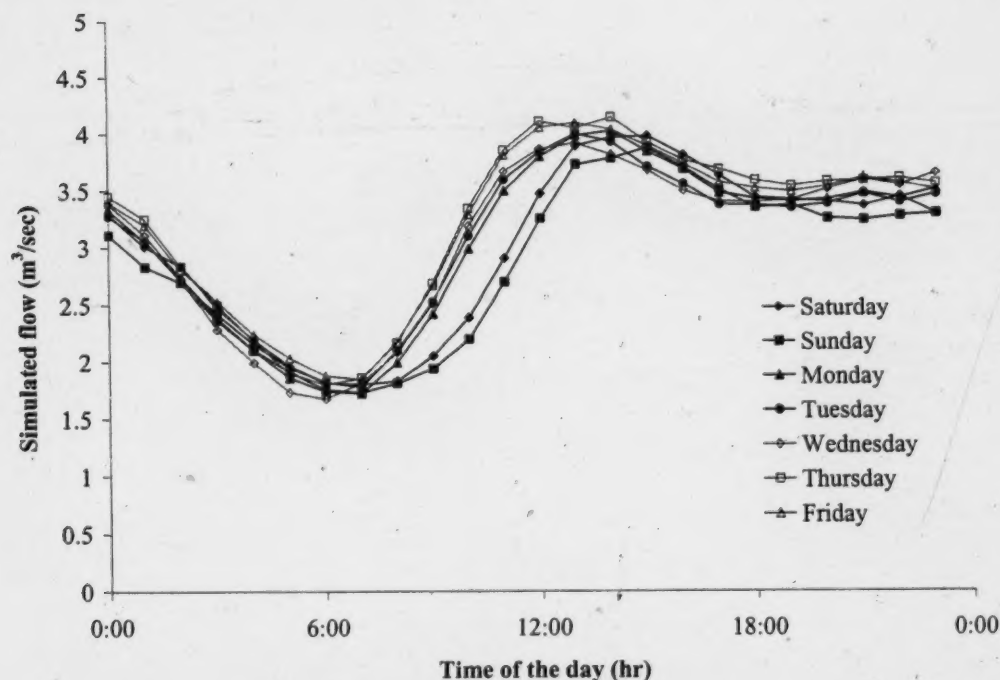


Figure 2.11 Dry weather flow patterns simulated by the neural network model.

It is evident that using the two indices of day-of-the-week and hour-of-the-day, the neural network was able to learn the diurnal changes of the dry weather flow at the entrance to the plant as well as the lag that existed between the weekend patterns and the weekdays ones. As was mentioned before, the main objective was to build a neural network that, when trained, would be able to generalize well to the validation data set that has never been seen by the network during the course of training. Model predictions for data set #1 that was used in training are shown in Figure 2.12 to Figure 2.19. The predictions shown in Figure 2.12 to Figure 2.19 were obtained when data set #3 was used as part of the training data set while data set #2 was used for validating the neural network. Figure 2.20 and Figure 2.21 show the model predictions when applied to the validation data. Data set #2 is shown in Figure 2.20 while data set #3 is shown in Figure 2.21. Figure 2.22 shows the prediction errors for the validation data.

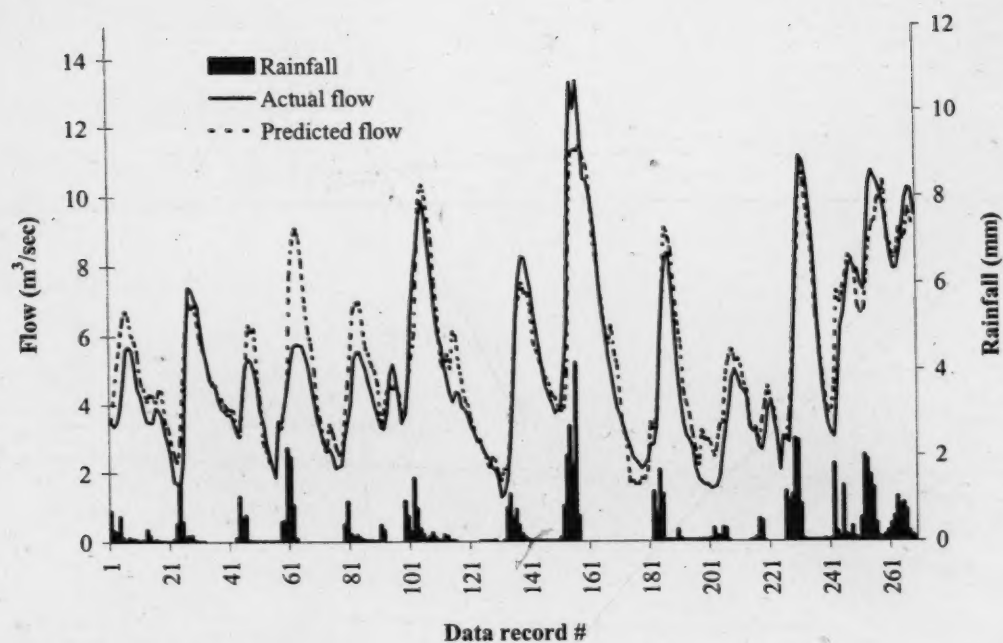


Figure 2.12 Actual versus predicted flow values for data set #1 - Part 1.

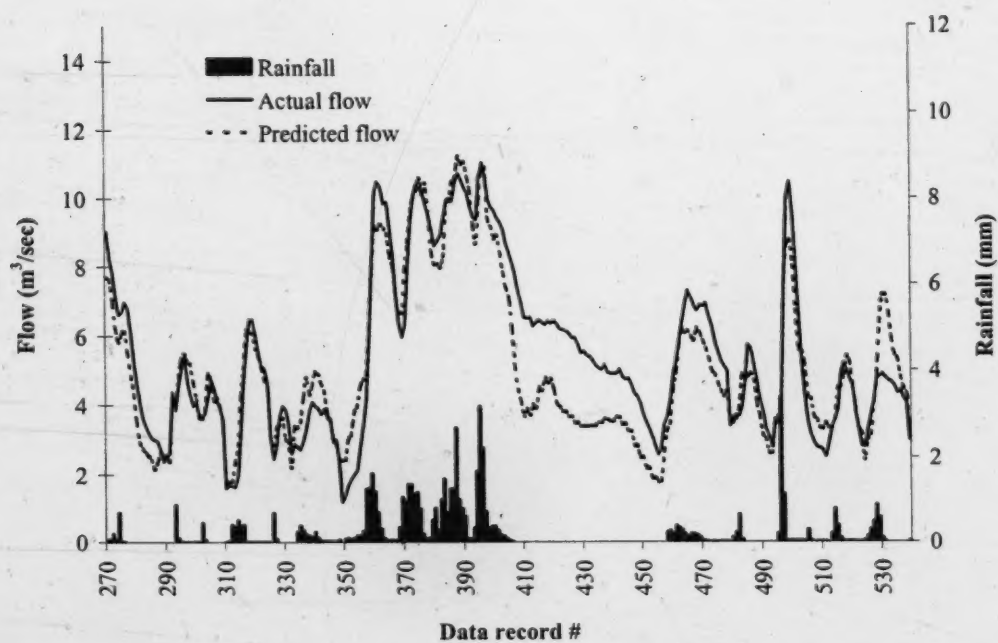


Figure 2.13 Actual versus predicted flow values for data set #1 - Part 2.

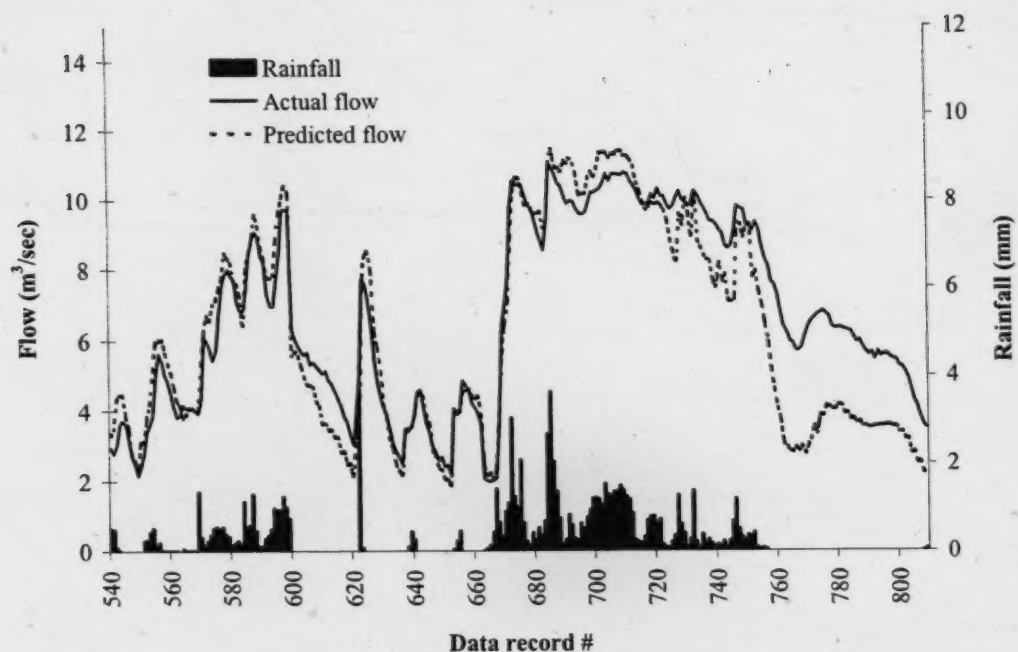


Figure 2.14 Actual versus predicted flow values for data set #1 - Part 3.

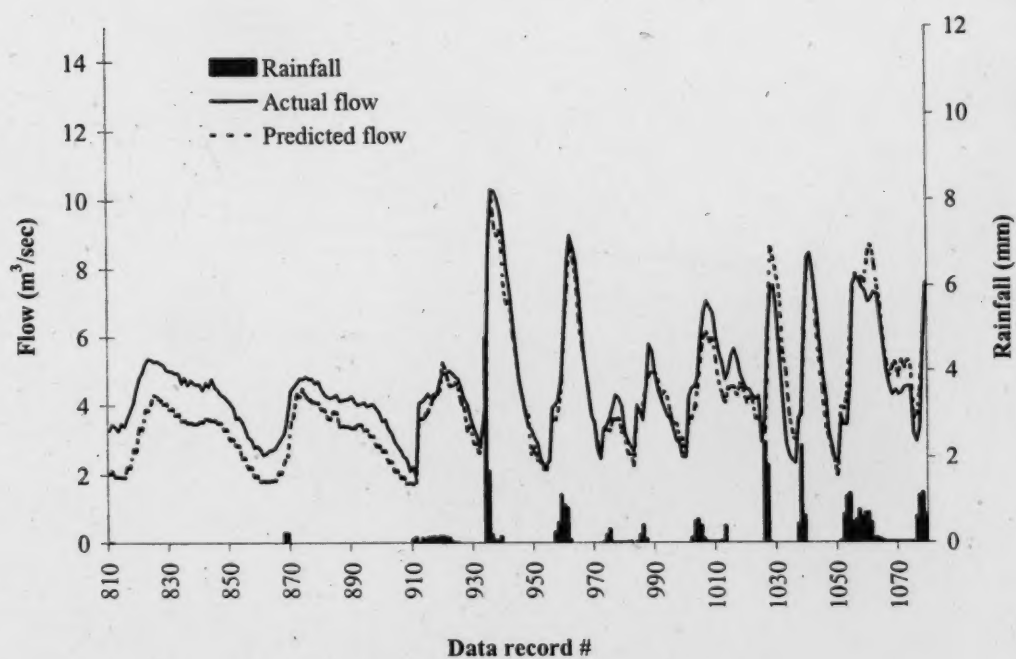


Figure 2.15 Actual versus predicted flow values for data set #1 - Part 4.

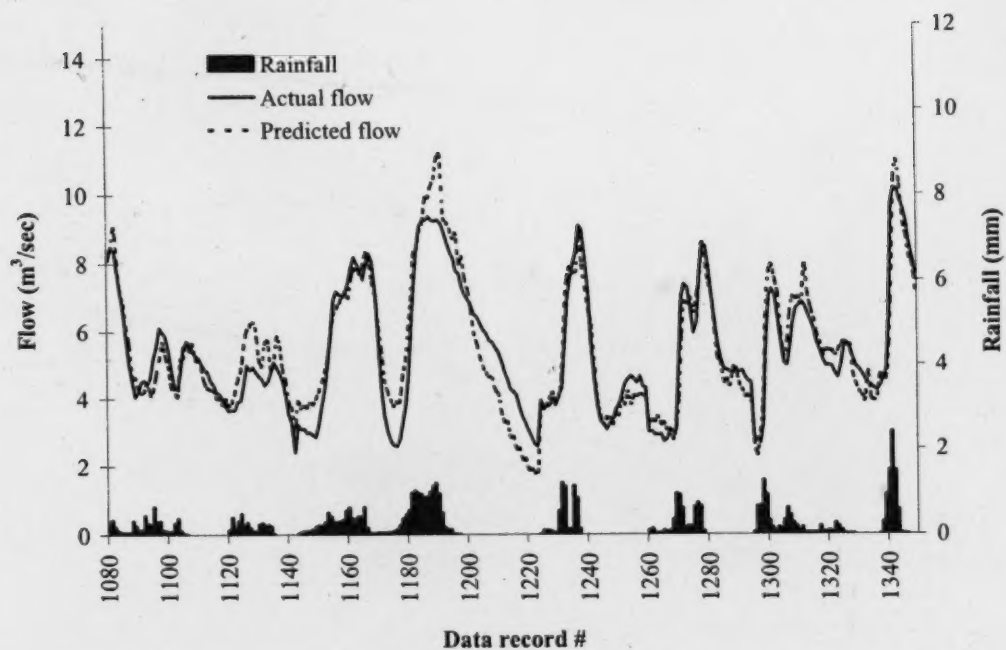


Figure 2.16 Actual versus predicted flow values for data set #1 - Part 5.

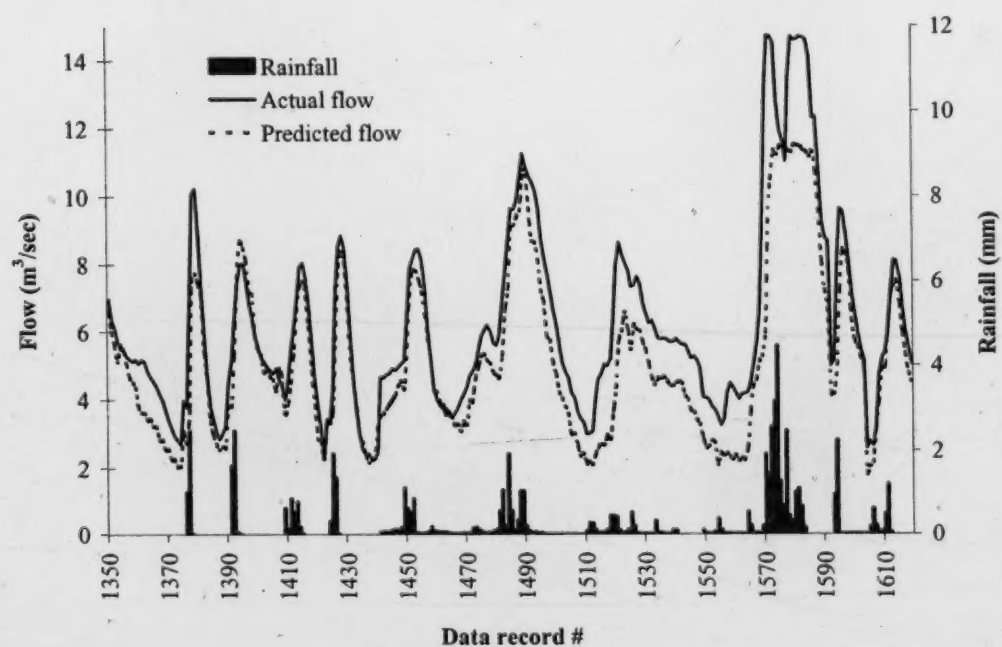


Figure 2.17 Actual versus predicted flow values for data set #1 - Part 6.

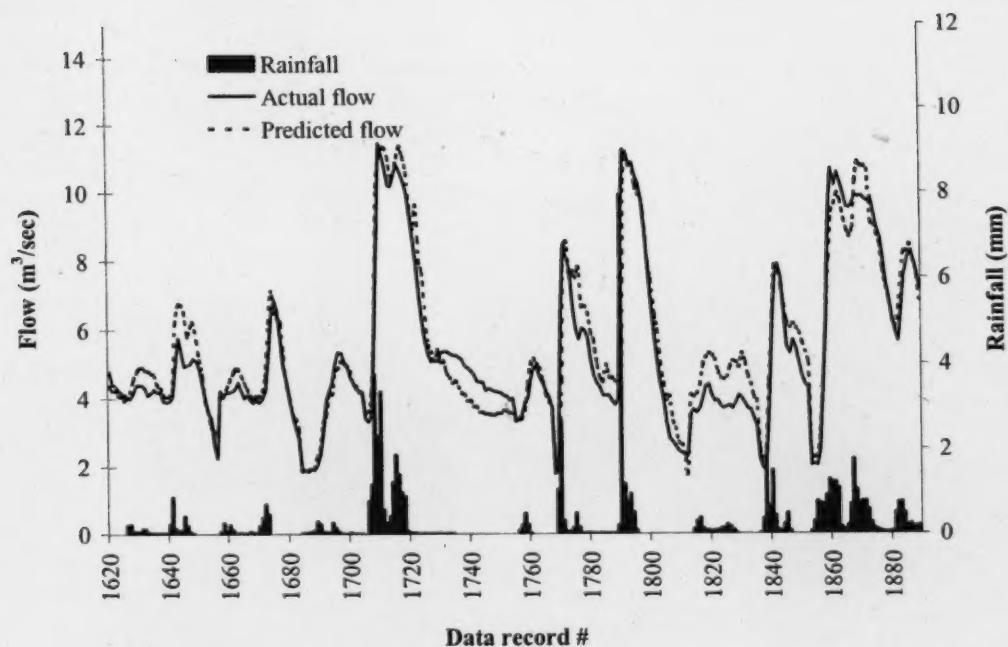


Figure 2.18 Actual versus predicted flow values for data set #1 - Part 7.

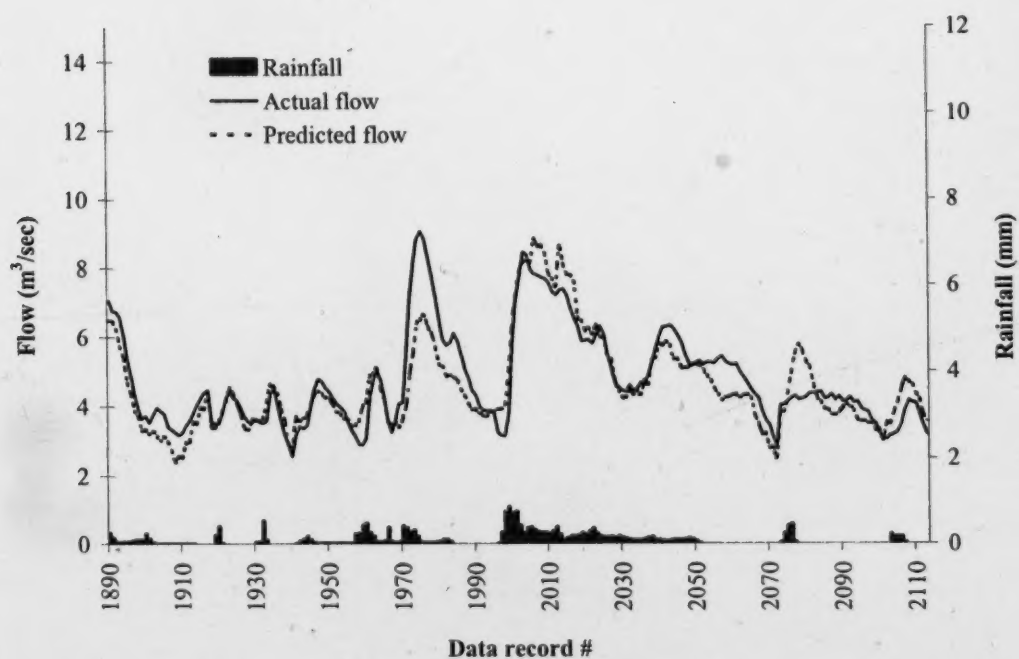


Figure 2.19 Actual versus predicted flow values for data set #1 - Part 8.

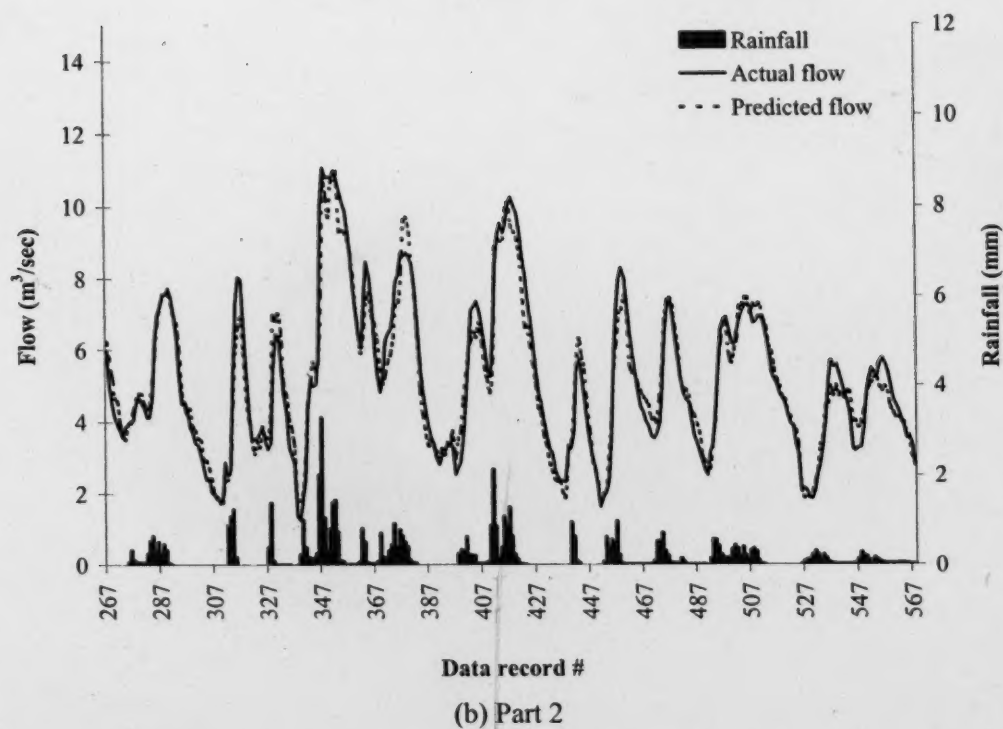
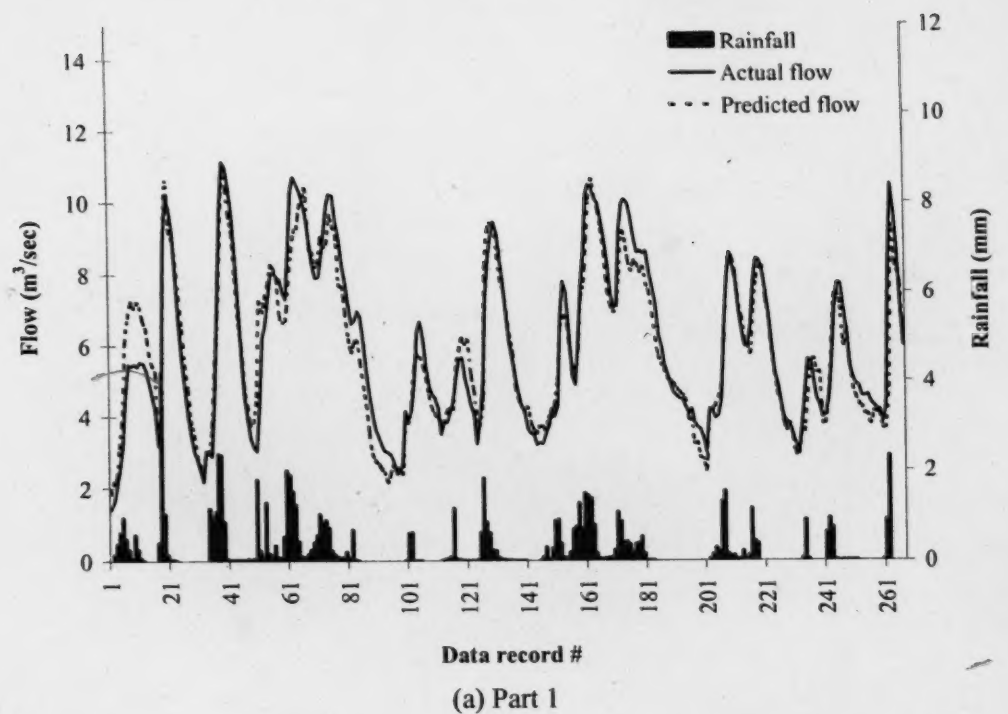
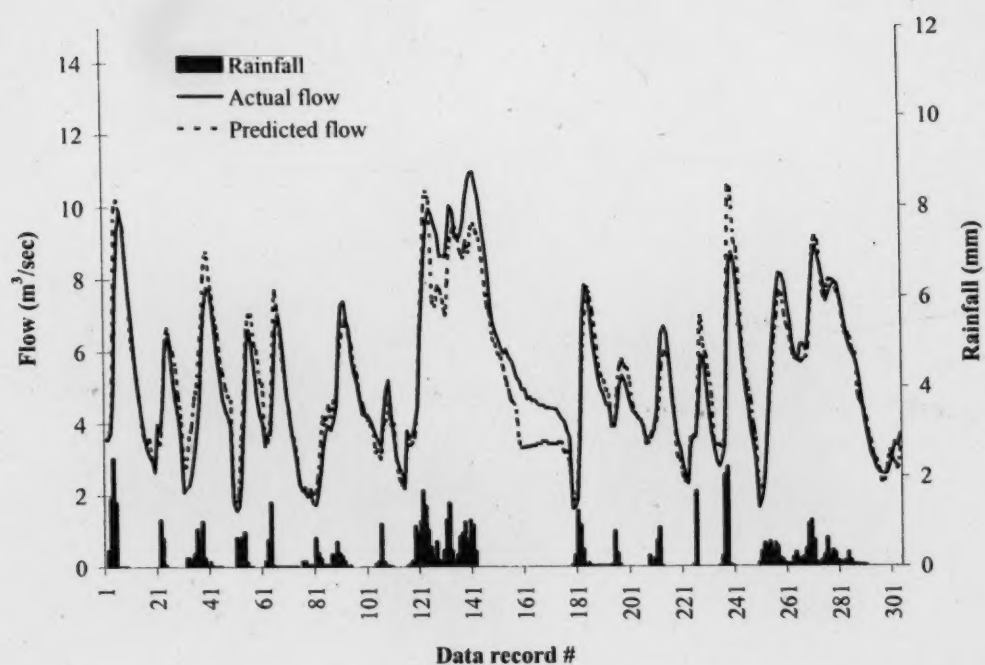
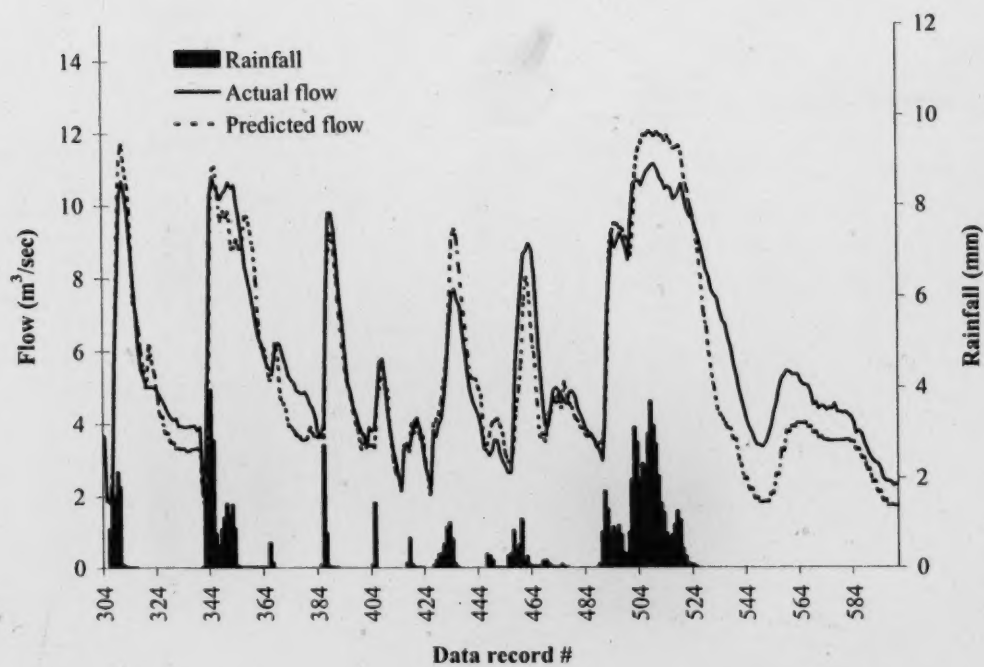


Figure 2.20 Actual versus predicted flow values for validation data set #2.

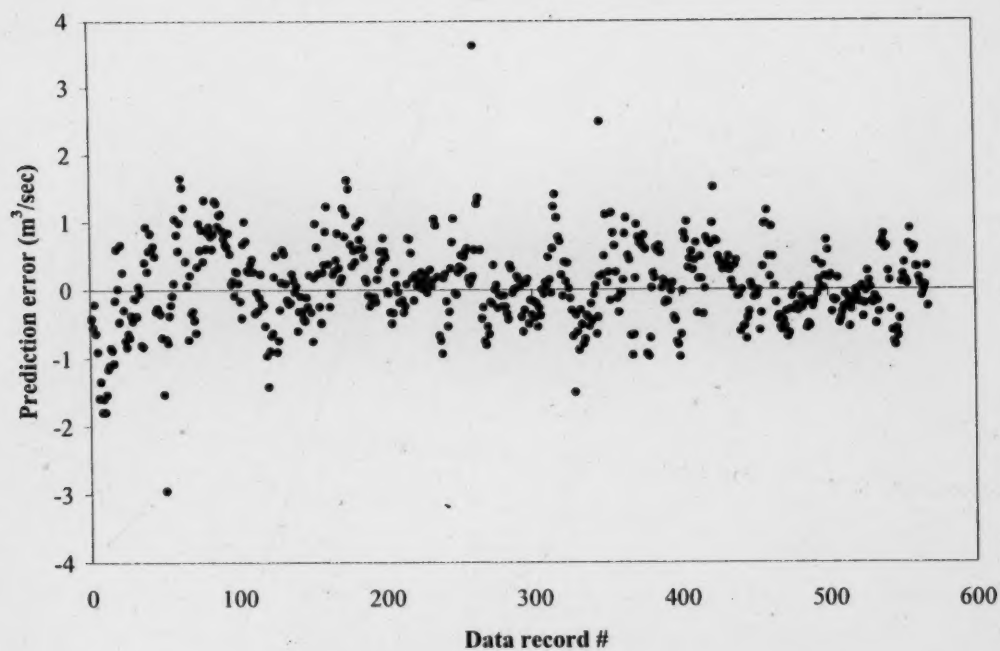


(a) Part 1

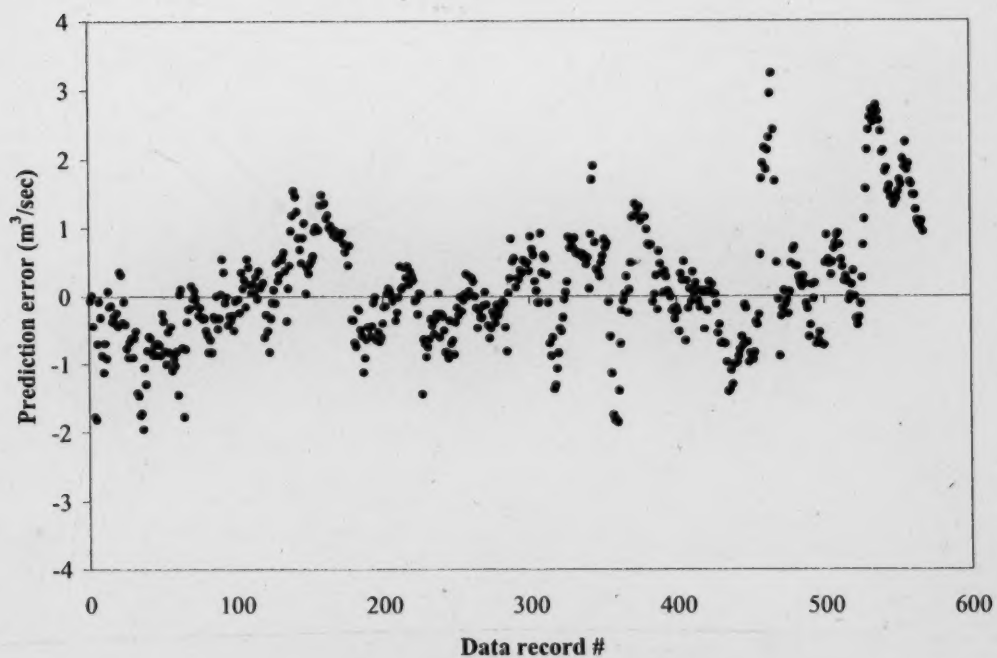


(b) Part 2

Figure 2.21 Actual versus predicted flow values for validation data set #3.



(a) when data set #2 was used for validation



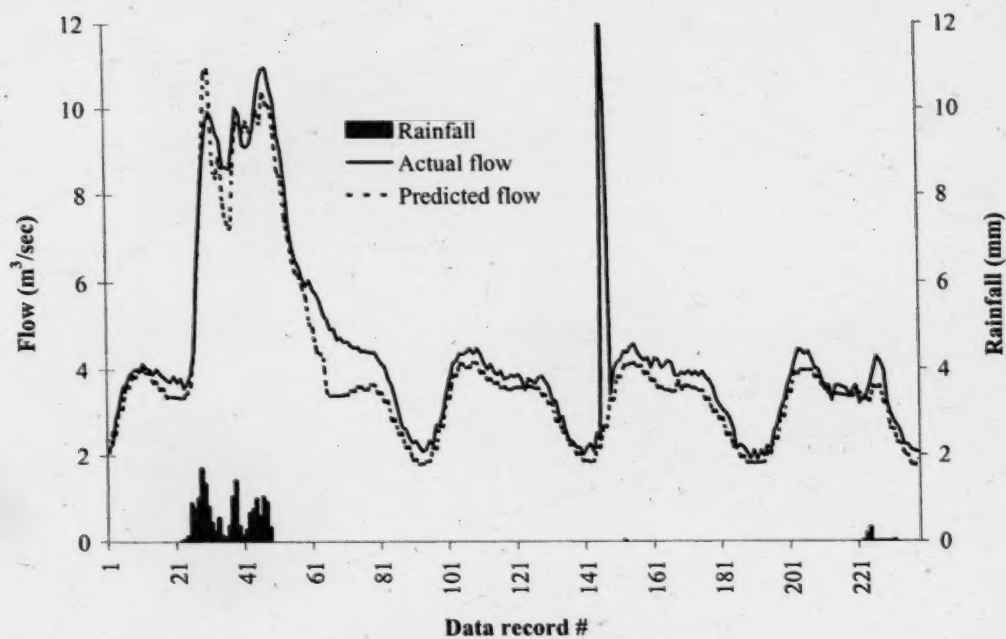
(b) when data set #3 was used for validation

Figure 2.22 Prediction errors for the validation data.

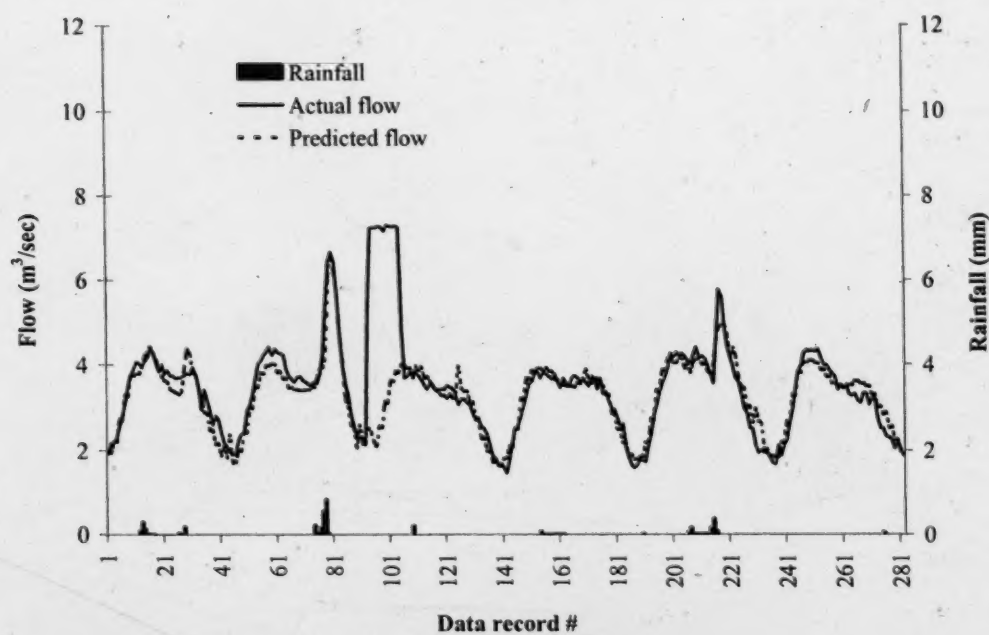
For both the training and validation data sets, the neural network model performed extremely well. When tested against the validation data that had not been seen by the model during the course of training, the model had excellent generalization ability and was able to predict fairly well the quantity of wastewater flow entering the plant for 44 different rainfall events that ranged widely in both the total amount of rainfall and the maximum intensity.

As was mentioned before, periods of faulty flow data had been visually identified and eliminated from the data used by the neural network model. However Figure 2.23 shows some periods of apparently erroneous flow data that are substantially above dry weather values without any corresponding rainfall data. These data records were left in the training data set in order to see their effect on the model behavior after being trained on a data set that are mostly of good quality except for very few cases of erroneous records. As evident from Figure 2.23, the model was not confused by these rare occasions which means that it can be updated (or re-trained) online without visual examination of training data that are presented to the model as long as the overall quality of the data is good.

One discrepancy existed between the model predictions and the actual flow data which is clear from Figure 2.21 (a) (Data records #160 to 180) and Figure 2.21 (b) (Data records #530 to 580). In each of these two cases, a rainfall event lasted more than 12 hours, and long after the event has stopped, the flow entering the plant was still above the normal dry weather flow values. This is because of the delayed inflow that took more than 8 hours (the size of the moving window) to reach the plant. As a result, the neural network model under-estimated the actual flows recorded at the entrance to the treatment plant for these data records. However, from a process-control point of view, this discrepancy is minor as long as the flow peaks are well predicted. Figure 2.20 to Figure 2.22 clearly indicate that the model did not have any obvious tendency to either over-estimate or under-estimate the flow peaks and there was no obvious lag between the actual peak recorded during a rainfall event and the one predicted by the neural network model.



(a) A spike of extremely high flow



(b) Consecutive high flow values

Figure 2.23 A portion of the training data set that contained erroneous flow data.

2.6 CONCLUSIONS

Wastewater flows at the GBWWTP can increase rapidly during rainfall events. The existing maximum capacity of the plant can only be achieved if warning is available that increased flows are expected. Therefore, a model that can predict the quantity of wastewater flow entering the plant during rainfall events is very beneficial for the operation of different unit processes during these events. The neural network model presented in this study can be utilized online to give short-term predictions of wastewater flow entering the plant. The model uses rainfall data recorded by 8 of the gauges that cover the City of Edmonton drainage area. A systematic approach was used in building the neural network model, which enabled it to give R^2 values in the range of 0.85 to 0.93 when it was tested against the validation data sets that were never seen before by the network during the course of training. The model can be integrated into a real-time control system with the ultimate goal of minimizing the total pollution load from the wastewater treatment system.

2.7 REFERENCES

1. Capodaglio, A.G., "Evaluation of Modelling Techniques for Wastewater Treatment Plant Automation", *Wat. Sci. Tech.* 30(2): 149-156 (1994).
2. Carstensen, J., M.K., Nielsen, and H. Strandbaek, "Prediction of Hydraulic Load for Urban Storm Control of a Municipal WWT Plant", *Wat. Sci. Tech.* 37(12): 363-370 (1998).
3. Cote', M., P.A. Grandjean, P. Lessard, and J. Thibault, "Dynamic Modelling of the Activated Sludge Process: Improving Prediction Using Neural Networks", *Wat. Res.* 29: 995-1004 (1995).
4. Djebbar, Y. and P.T. Kadota, "Estimating Sanitary Flows Using Neural Networks", *proc. Of IAWQ 19th Biennial International Conference.*: Vancouver, Canada, 1998.
5. Gallant, S.I., *Neural Network Learning and Expert Systems* (Cambridge, MA, USA: MIT Press, 1993).
6. Lippmann, R.P., "An Introduction to Computing with Neural Nets", *IEEE ASSP Magazine.* April: 4-22 (1987).
7. Rumelhart, D.E., G.E. Hinton, and J.L. McClelland, *A General Framework for Parallel Distributed Processing. Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (Cambridge, MA, USA: MIT Press, 1: 45-76, 1986).
8. Rumelhart, D.E., G.E. Hinton, and R.J. Williams, *Learning Internal Representations by Error Propagation. Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (Cambridge, MA: MIT Press, 1: 318-362, 1986).
9. Tyagi, R.D. and Y.G. Du, "Kinetic Model for the Effects of Heavy Metals on Activated Sludge Process Using Neural Networks", *Env. Tech.* 13: 883-890 (1992).
10. Zhao, H., O.J. Hao, T.J. McAvoy, and C.H. Chang, "Modelling Nutrient Dynamics in Sequencing Batch Reactor", *J. Env. Eng., ASCE* 123(4): 311-319 (1997).

CHAPTER 3. MODELLING A FULL-SCALE PRIMARY SEDIMENTATION TANK USING ARTIFICIAL NEURAL NETWORKS*

3.1 INTRODUCTION

Removing readily settleable solids and floating material and thus reducing the suspended solids content is the objective of primary treatment. Primary sedimentation tanks generally provide suspended solids (SS) and biochemical oxygen demand (BOD) reduction before further processing of the wastewater by secondary treatment. Modelling the performance of full-scale primary sedimentation tanks has been commonly done using regression-based models. These models are empirical relationships that are derived strictly from observed daily average influent and effluent data. A relationship of this type has been proposed by Smith (1969) based on data collected at a large number of different plants with different flow rates. The equation proposed by Smith (1969) relates the total suspended solids removal to the overflow rate in a nonlinear fashion suggesting that the performance of a sedimentation tank will always deteriorate at higher flow rates. Using only the overflow rate to model the total suspended solids removal by a sedimentation basin is basically derived from the theoretical consideration of the process.

Tebbutt (1969) conducted pilot-scale studies over a wide range of hydraulic loadings and showed a lack of correlation between the overflow rate and the settling performance of sedimentation tanks and he attributed these findings to the flocculent nature of wastewater solids which can cause the formation of readily settleable solids in the presence of the type of velocity gradients created by turbulence and density currents. Since flocculation will be more pronounced at higher suspended solids concentrations, other workers (Steel, 1960; Voshel and Sak, 1968; Tebbutt and Christoulas, 1975) have suggested that both the hydraulic loading and the influent suspended solids concentration affect the removal efficiency of a sedimentation tank. In addition, the characteristic of the solids in terms of their ability to stick together upon impact is also important. Therefore, a tank treating an influent that is low in suspended solids with relatively low removal might be more efficient at removing settleable solids when it is given a stronger influent.

A version of this chapter has been published. Gamal El-Din, A. and D.W. Smith, *Env. Tech.* 23: 479-496 (2002).

A mathematical model of sedimentation performance that has been commonly used is of the general form (Tebbutt and Christoulas, 1975):

$$E = a \cdot e^{-(b/S+cq)} \quad [3.1]$$

where

E = the fractional removal of suspended solids,

S = the influent suspended solids,

q = the surface overflow rate, and

a , b , and c = constants.

It is clear from the above equation that at the same hydraulic loadings, when wastewater influent that is rich in total suspended solids is introduced to a sedimentation tank, the removal efficiency of the tank will increase due to the more pronounced flocculation that will occur within the tank.

The primary role of sedimentation units in a conventional treatment plant is the removal of suspended matter, however in terms of the effect of sedimentation on the downstream biological processes the removal of organic matter in the primary stage is of considerable importance (Tebbutt and Christoulas, 1975). The organic content of domestic wastewater as measured by chemical oxygen demand (COD) is present in three forms, large suspended matter, colloidal solids, and soluble compounds. All of the large solids and some of the fine material are potentially removable by sedimentation, however, most of the fine material and soluble organics are not removable to any significant extent. The removal of organic matter by primary sedimentation has been commonly done using empirical models that utilize daily average data and correlate the removal of organic matter to the removal of suspended matter.

3.2 DYNAMIC MODELLING

A wastewater treatment plant usually has a satisfactory performance under steady-state conditions because these conditions are similar to design conditions. However, load variations constitute a large portion of the operating life of a treatment facility and most of the observed problems in complying with permit requirements are due to these load transients. Inflow to wastewater treatment facilities is a random variable that necessarily affects the performance of the facility. The resulting transient performance as well as the need to characterize the impacts of this variability on the removal efficiency of the facility has been widely recognized (Alarie et al., 1980). Although several levels of sewage treatment are practiced throughout North America, primary sedimentation is a unit treatment process common to the wastewater treatment facilities for most medium to large cities, and therefore, improvements in the operation of a sedimentation tank, the first major unit process in the train of processes, are likely to produce a pronounced impact on the quality of overall treatment (Alarie et al., 1980).

Hydraulic efficiency models have been used to characterize the dynamic behavior of sedimentation tanks (El-Baroudi, 1969; Thirumurthi, 1969). These models recognize that real flows are turbulent and encounter a certain degree of mixing or eddy diffusion. The developers of these models utilize tracer studies to characterize the performance of model sedimentation tanks based on eddy diffusion. Using these models to predict the dynamic response of a full-scale sedimentation tank is very difficult as the development of such models has been done using controlled studies of model tanks (Alarie et al., 1980). In order to model the dynamic behavior of a sedimentation tank under transient conditions, some deterministic models are available in the literature (Takamatsu et al., 1974; Shiba and Inoue, 1975) that are based on a mathematical formulation that takes into consideration the physical configuration of a sedimentation tank and depicts the physical phenomena of deposition and resuspension occurring within the tank. Although these models permit detailed simulation of primary sedimentation performance, they were theoretically derived and have not been tested with actual field data.

Another type of model is the so-called "black-box" class of models. As the name implies, these models do not try to describe the internal functions of a system, however, they can recognize relationships that exist between measurable inputs and outputs of the system. Among these models, are those using the univariate and multi-variate Box-Jenkins modelling approach and artificial neural networks (ANNs). The main advantage of neural network models over a parametric approach is that neural networks can generate complicated relationships through examination of only the data points in the training set without assuming a pre-specified functional form (Djebbar and Kadota, 1998). Therefore, neural networks have been successfully applied in many areas in the environmental engineering field. The objective of this study was to model the performance of a full-scale primary sedimentation tank at the Gold Bar Wastewater Treatment Plant (GBWWTP) under dynamic loading conditions using an artificial neural network model. An overview of ANN is provided by Gamal El-Din and Smith (2002).

3.3 DESCRIPTION OF THE TREATMENT PLANT

The Gold Bar Wastewater Treatment Plant (GBWWTP) was constructed in 1956 on the south west shore of the North Saskatchewan River. The present capacity of the plant is 950 (ML/d) for primary treatment and 420 (ML/d) for secondary treatment. The plant treats domestic and industrial sewage from the City of Edmonton. The GBWWTP is typical of many conventional activated sludge plant designed for carbonaceous BOD and suspended solids removal. It provides both primary and secondary treatment for the incoming raw sewage. Primary treatment consists of a raw influent distribution chamber, two Venturi flume installations, five aerated grit chambers, six bar screens and eight rectangular primary settling tanks. The secondary treatment provides biological treatment in a suspended growth activated sludge system, final settling, and microorganism reduction.

Table 3.1 The physical parameters of primary settling tanks group 2 (PST 2).

Primary Settling Tanks Group 2 (PST 2)	
No. of tanks	4
Tanks #	#5, 6, 7, and 8
Length (m)	51
Width (m)	20
Total volume (m ³)	19,000
Total surface area (m ²)	4,160
Weir length (m)	302
Peak design loading per tank (ML/d)	112

There are two primary sections in the plant, Primary Settling Tanks Group 1 (PST 1), which includes settling tanks #1, 2, 3, and 4, and Primary Settling Tanks Group 2 (PST 2) which includes settling tanks #5, 6, 7, and 8. The distribution of the incoming wastewater flow between the two sections is controlled by two manually activated sluice gates located upstream of the Venturi flume installation and by the difference in throat width between the two flumes.

Downstream from the bar screens the effluent flows into four channels, each controlled by a sluice gate, and feeding wastewater to two primary tanks. These gates are operated as isolation gates, and are either fully opened or fully closed. The flow into each tank is controlled by opening or closing the inlet ports. The arrangement of primary tanks on common influent channels is: #1 and 2; #3 and 4; #5 and 6; and #7 and 8. All eight primary sedimentation tanks operate in a similar manner. Each tank is divided into four cells. Each cell has two inlet ports, a sludge and scum collection system, and two effluent weirs. Primary sedimentation tank #5 (one of the four tanks of PST 2) was chosen for sampling. Table 3.1 shows the physical parameters of tanks of PST 2.

3.4 SAMPLING PROGRAM

The influent wastewater to the tank was sampled at the common channel that feeds both tanks #5 and 6. The effluent wastewater from the tank was sampled from the first trough at the first cell from the left of the tank as access of sampling the effluent from the sedimentation tank was only available at this point. The flow of wastewater received by PST 2, measured by the Venturi flume installation, was averaged over 30-minute intervals. Two weeks of continuous sampling were conducted. The first week started on June 28, 1999 at 6:00 hours and ended July 5, 1999 at 6:00 hours. The second week started on August 20, 1999 at 6:00 hours and ended August 27, 1999 at 6:00 hours.

For both periods, grab samples for the total suspended solids (TSS) analysis were taken every hour, from both the influent and effluent. At 7:00 hours of every day during the sampling period, the TSS samples of the previous 24-hour were taken to the laboratory at the University of Alberta for immediate analysis. From July 2, 6:00 hours till July 5, 6:00 hours and from August 20, 6:00 hours till August 27, 6:00 hours, grab samples for the chemical oxygen demand (COD) analysis were taken every hour, from both the influent and effluent and were acidified by concentrated sulfuric acid to pH less than 2 for the purpose of preservation. At 7:00 of every day, the COD samples of the previous 24-hour were taken to the laboratory at the University of Alberta and stored in the cold room at 4 °C to be analyzed later. Both air and influent wastewater temperatures were recorded every hour.

3.5 METHODS

The analysis of total suspended solids was according to the Standard Methods (APHA, 1995). For both influent and effluent TSS samples, the analysis was done in three replicates for each sample and the average was utilized. Before the three replicates were drawn off a sample, blending the sample for one minute was done in order to have representative samples. Whatman grade 934AH glass micro-fiber filters were used (Whatman catalog #1827 024). The filters have a diameter of 24 mm and a pore size of

1.50 μm . In practice, it is unlikely that particles smaller than 1.50 μm would be removable by primary sedimentation. Tebbutt (1979) stated that particles much smaller than 50 μm would be unlikely removed by sedimentation. Therefore, using a filter of a pore size 1.5 μm was thought to be appropriate for the modelling of the sedimentation tank and, at the same time, facilitated the analysis of the large number of samples (144 samples) that had to be analyzed daily utilizing all the analytical support staff that was available.

The analysis of chemical oxygen demand was done according to the Standard Methods (APHA, 1995) using the closed reflux colorimetric method. For effluent COD samples, the analysis was done in three replicates for each sample and the average was used. For influent COD samples, the analysis was done in five replicates and the average was utilized. Before the five replicates were drawn off an influent sample, blending the sample for one minute was done in order to have representative samples.

3.6 RESULTS OF THE SAMPLING PROGRAM

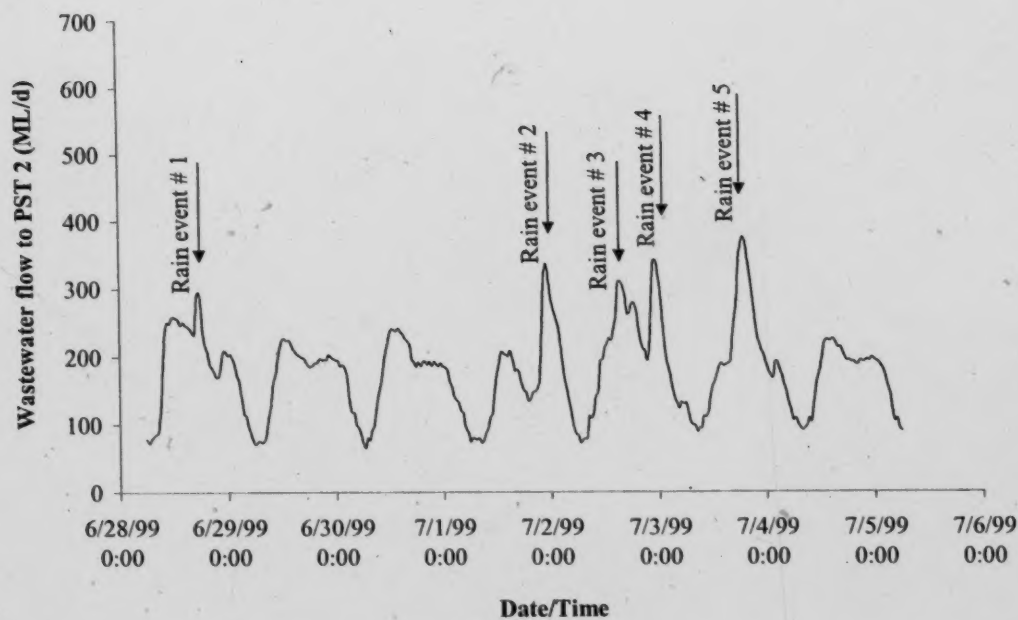
Descriptive statistics of the data sets that were collected during the sampling periods are shown in Table 3.2. The wastewater inflow to PST 2 recorded during the first and second weeks of sampling is shown in Figure 3.1 (a) and (b), respectively. During the sampling periods, few rainfall events occurred which differed in intensity as well as duration. During the first week of sampling five rainfall events occurred and are indicated in Figure 3.1 (a). Only one rainfall event occurred during the second week of sampling and is indicated in Figure 3.1 (b). Event #6, which commenced on August 21, 17:00, had the highest intensity and during this event the flow increased by more than 200% of the normal dry weather values. The second largest event was #5 which commenced on July 3, 15:00 and during which the increase in flow was recorded to be around 100% of the normal dry weather values. The flow data indicated that the normal dry weather flow values recorded in the days of August were higher by around 25% than those values of June and July.

Table 3.2 Descriptive statistics of the data.

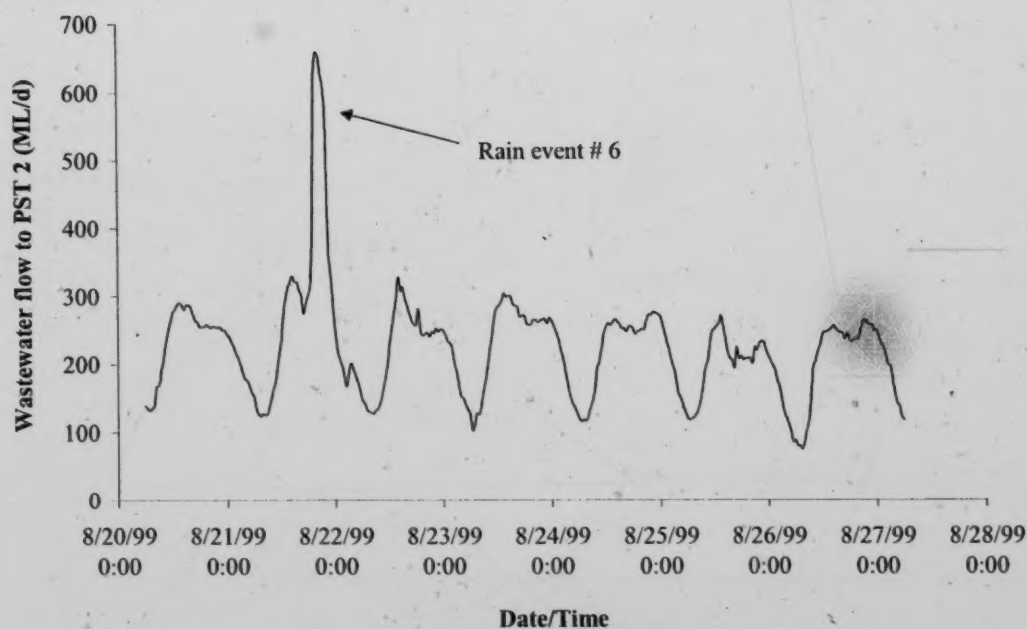
Sampling period	Parameter	Average (mg/L)	Standard deviation (mg/L)	Min. (mg/L)	Max. (mg/L)
June 28 – July 5, 99	Flow	179 ^a	65 ^a	66 ^a	375 ^a
June 28 – July 5, 99	Influent TSS	207	95	80	682
June 28 – July 5, 99	Effluent TSS	47	20	13	117
July 2 – July 5, 99	Influent COD	534	81	384	683
July 2 – July 5, 99	Effluent COD	358	59	237	477
August 20 – 27, 99	Flow	224 ^a	80 ^a	76 ^a	658 ^a
August 20 – 27, 99	Influent TSS	226	65	113	598
August 20 – 27, 99	Effluent TSS	48	24	13	239
August 20 – 27, 99	Influent COD	671	103	399	1058
August 20 – 27, 99	Effluent COD	450	71	281	626

^aML/d

From August 22, 6:00 hours till August 27, 6:00 hours no rainfall events occurred. This was the longest period of continuous dry weather flow conditions for these data sets. The flow profiles were averaged over the 5-day stretch in order to visualize a typical dry weather flow profile, which is shown in Figure 3.2 after being normalized with respect to the average. Between 2:00 and 11:00 hours, the flow is below average with a minimum value of 0.53 at 6:30 hours, and between 11:00 and 2:00 hours the flow is above average with a maximum value of 1.32 at 14:00 hours. During the early morning hours minimum flows occur when water consumption is lowest and when the base flow consists of infiltration and smaller quantities of sanitary wastewater. Two peak flows are evident from Figure 3.2. The first peak flow generally occurs in the afternoon, around 14:00 hours when the wastewater from the peak morning water use reaches the treatment plant. A second lower peak generally occurs in the evening, around 22:30 hours when the wastewater from the evening water use reaches the plant.



(a) First week of sampling



(b) Second week of sampling

Figure 3.1 Wastewater flow recorded during the sampling periods.

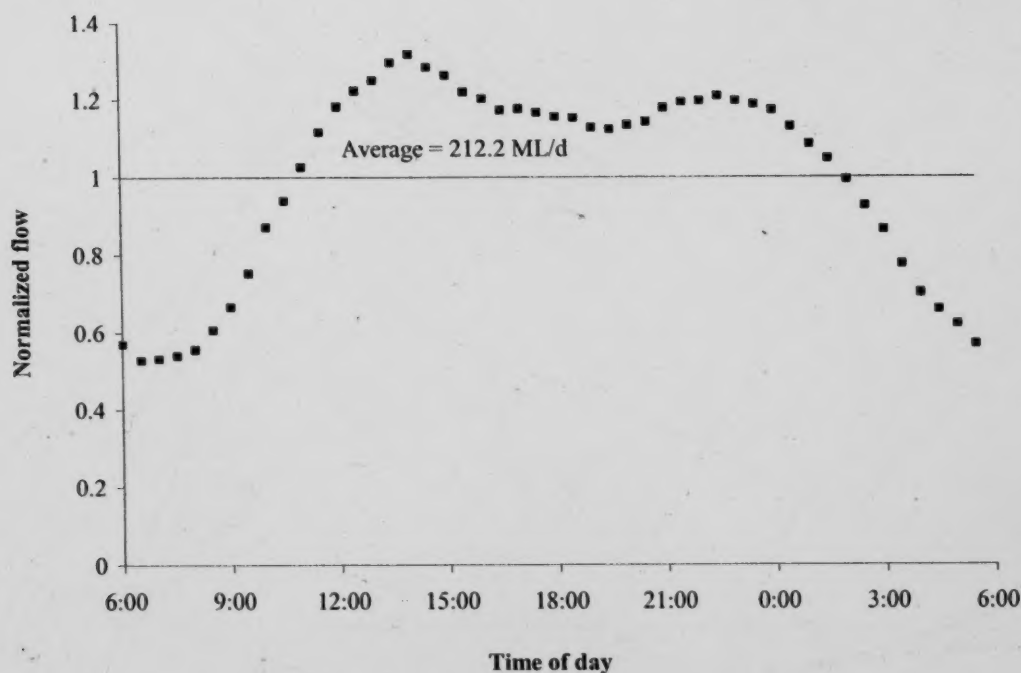
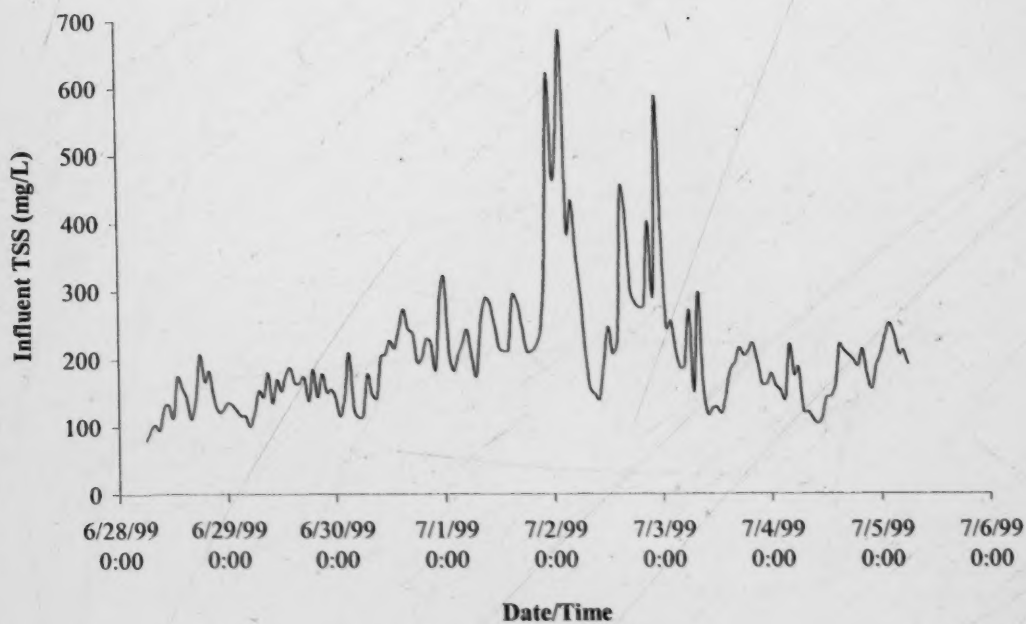


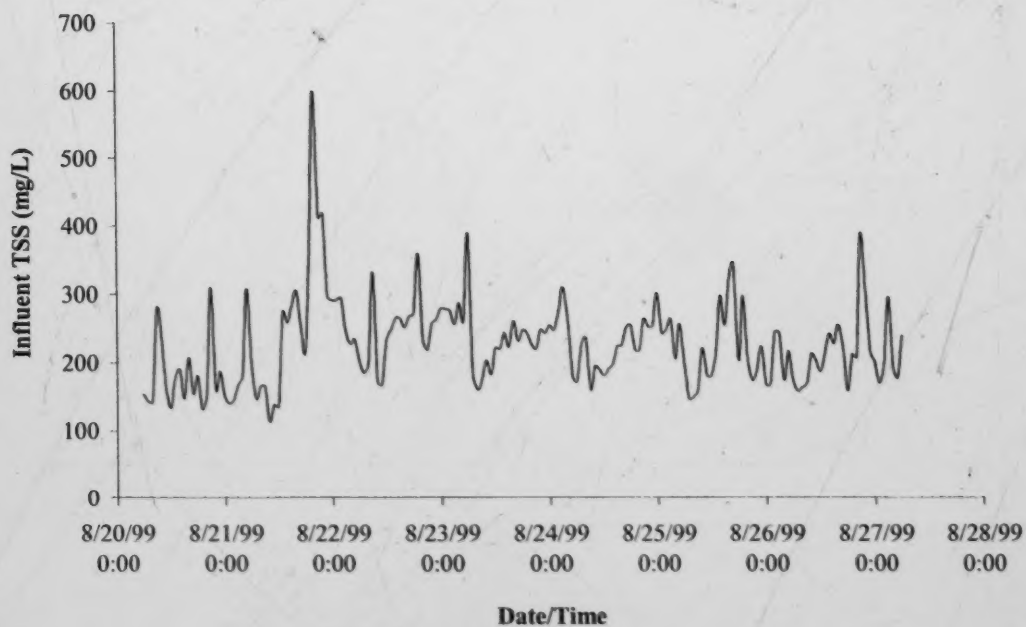
Figure 3.2 Dry weather flow profile.

3.6.1 Total Suspended Solids Data

Both the hydraulic and organic loadings to wastewater treatment plants have a fluctuating component with the fluctuations in quality of wastewater being more difficult to describe and quantify. Both long-term (seasonal, monthly or weekly) and short-term (diurnal, hourly, and even shorter frequency) fluctuations exist. Figure 3.3 (a) shows the TSS values measured for the influent samples collected during the first week of sampling. The influent TSS values for the second week of sampling are shown in Figure 3.3 (b). Table 3.2 shows that for the first week of sampling, the standard deviation of the influent TSS data was approximately 46% of the average, however for the second week of sampling it was only 29% of the average value and the reason for that is the different weather conditions experienced during each week.



(a) First week of sampling



(b) Second week of sampling

Figure 3.3 Influent total suspended solids.

During the first week, 5 rainfall events occurred, however during the second week, only one event occurred. During wet weather conditions the flow entering the plant is composed of wastewater and rainfall runoff that contributes to the system. Generally storm runoff will contribute a larger amount of flow which is of better quality than the wastewater, and therefore, COD and fecal coliform bacteria concentrations in the wastewater inflow are expected to be generally low during a rainfall event. On the other hand, suspended solids concentrations in the influent wastewater are normally elevated during a rainfall event, mainly due to the effect of the first flush. Factors that affect the magnitude of the first flush effect include the length of the period of dry weather flow that preceded the rainfall event, the combined sewer slopes (which influence the velocity of the flow), street and catch basin cleaning frequency and design, rainfall intensity and duration, and the frequency of cleaning the sewers (Metcalf & Eddy Inc., 1991). In the present study, the influent TSS increased substantially during most of the six rainfall events that were encountered during the sampling periods. These findings have been found elsewhere (Vanderborght and Wollast, 1990; Bertrand-Krajewski, 1992; Bertrand-Krajewski et al., 1995). Among the 6 events, rainfall event #6 had the highest flow contribution to the plant. At the time of the commencement of the event, the flow was around 275 ML/d and during the event, the flow increased rapidly to a value of 657 ML/d as shown in Figure 3.1 (b). This event caused the influent TSS to increase from values around 260 mg/L to a value of 600 mg/L, which means that during this event, the influent TSS increased by more than 120% of the dry weather values.

As was mentioned before, August 22, 6:00 hours to August 27, 6:00 hours was the longest period of continuous dry weather flow conditions. The influent TSS hourly profiles were averaged over the 5 day stretch in order to visualize a typical hourly influent TSS profile during dry weather conditions, which is shown in Figure 3.4 after being normalized with respect to the average. It is clear from Figure 3.4 that a diurnal pattern exists, which is less evident than the one for the wastewater inflow. Between 5:00 and 13:30 hours the influent TSS is lower than average with a minimum value of approximately 0.8 of the average at around 9:00 hours. The influent TSS increases

gradually from 9:00 till around 17:00 hours then stays slightly above average till around 5:00 hours.

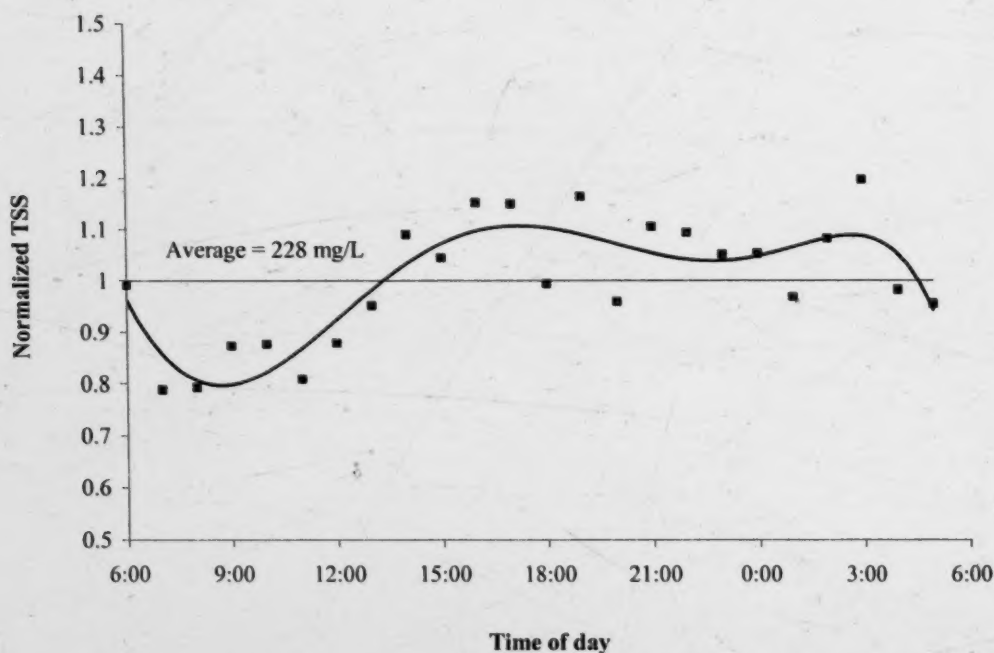


Figure 3.4 Influent total suspended solids profile during dry weather conditions - August 22 to 27, 1999.

3.6.2 Chemical Oxygen Demand Data

Figure 3.5 shows the COD values measured for the influent samples collected for the 10 days period of sampling for COD. The influent COD hourly profiles were averaged over the 5 day stretch of August 22, 6:00 to August 27, 6:00 hours, during which no rainfall events occurred, in order to visualize a dry weather profile for the influent COD which is shown in Figure 3.6 after being normalized with respect to the average. A diurnal pattern exists which is more evident than the one for the influent TSS. Between 6:00 and 14:00 hours the influent COD is lower than average with a minimum value of approximately 0.85 of the average base line at around 10:00 hours. The influent COD increases gradually from 10:00 till around 18:00 hours when it reaches a peak value

of 1.10 then decreases gradually from 18:00 till 23:00 hours then stays slightly above average till around 5:00 hours.

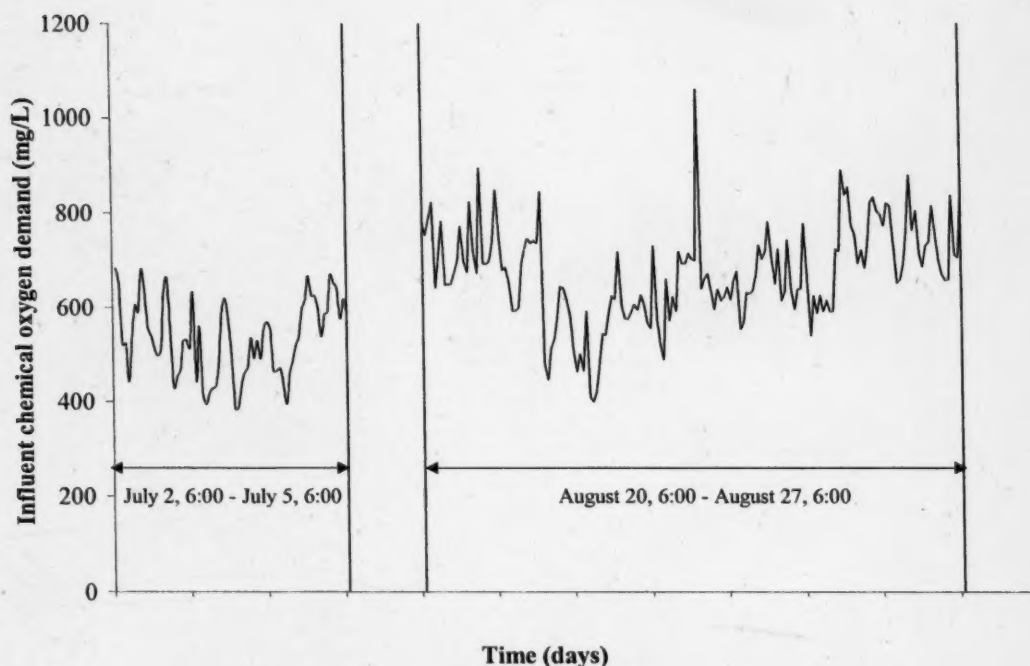


Figure 3.5 Influent chemical oxygen demand.

During the period of COD sampling, 4 rainfall events occurred, namely events #3, 4, 5, and 6. Each of these events had a pronounced effect on the influent COD values measured, which dropped substantially during each of the rainfall events encountered due to the dilution effect of the storm water, and then after each of these storms, when the combined flow was mainly wastewater, concentrations increased. This phenomenon was most evident for event #6 which had the most contribution of storm runoff to the sewer system as indicated by the flow records during the period of this rainfall event. During this event, the influent COD values had a steep drop from approximately 730 mg/L to around 450 mg/L then increased gradually after the end of the event.

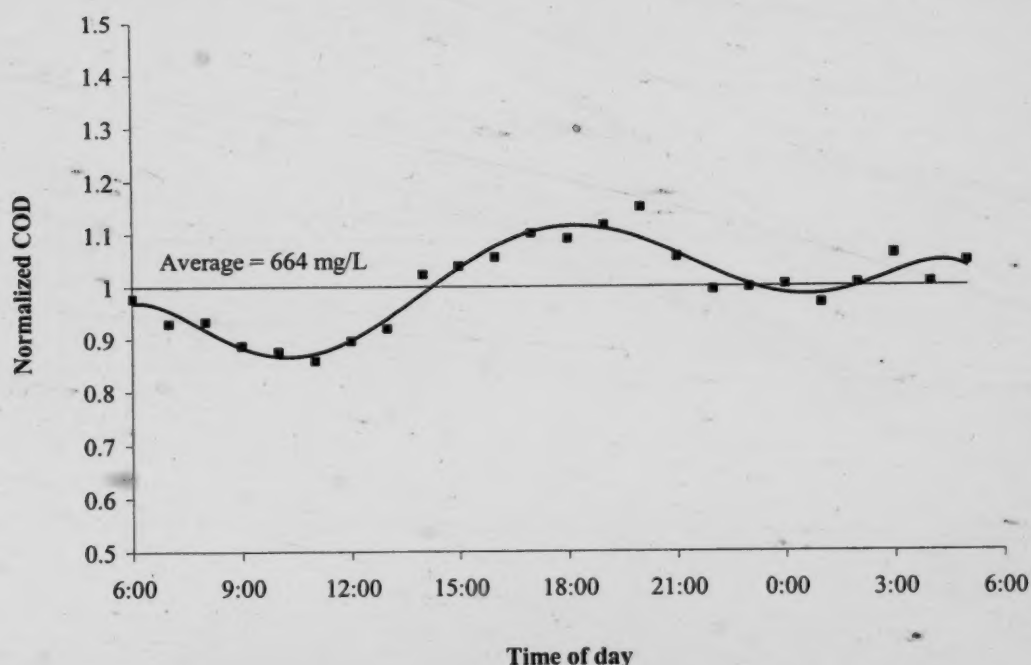


Figure 3.6 Influent chemical oxygen demand profile during dry weather conditions - August 22 to 27, 1999.

3.7 MODELLING THE EFFLUENT TOTAL SUSPENDED SOLIDS AND CHEMICAL OXYGEN DEMAND

3.7.1 Theoretical Detention Time

The total volume of the four sedimentation tanks in PST 2 (tanks #5, 6, 7, and 8) is equal to 19,000 m³. During the two week sampling periods, the flow entering the four tanks averaged 202 ML/d. The minimum flow was 66 ML/d and the maximum was 658 ML/d. Assuming equal flow distribution between the four tanks, the theoretical detention time for each tank is shown in Figure 3.7. The four sedimentation tanks in PST 2 have been designed based on an average flow to the process of 204 ML/d which corresponds to a theoretical detention time of 2.24 hr. This design value is very close to the average flow that entered PST 2 during the sampling periods. It is clear from Figure 3.7 that the minimum flow received by PST 2 corresponds to a theoretical detention time of 6.91 hr while the maximum flow corresponds to a 0.69 hr theoretical detention time. It should be

mentioned that there is a delay between the time of recording the flow and the time of the wastewater entering the sedimentation tank. This delay is equal to the residence time of the wastewater in the aerated grit chambers upstream from the primary settling tanks of PST 2. Based on the average flow of 200 ML/d received by PST 2, and the theoretical detention time of the grit chambers, this delay is approximately 15 minutes which is 25% of the sampling frequency (one hour), and as a result, this discrepancy was ignored.

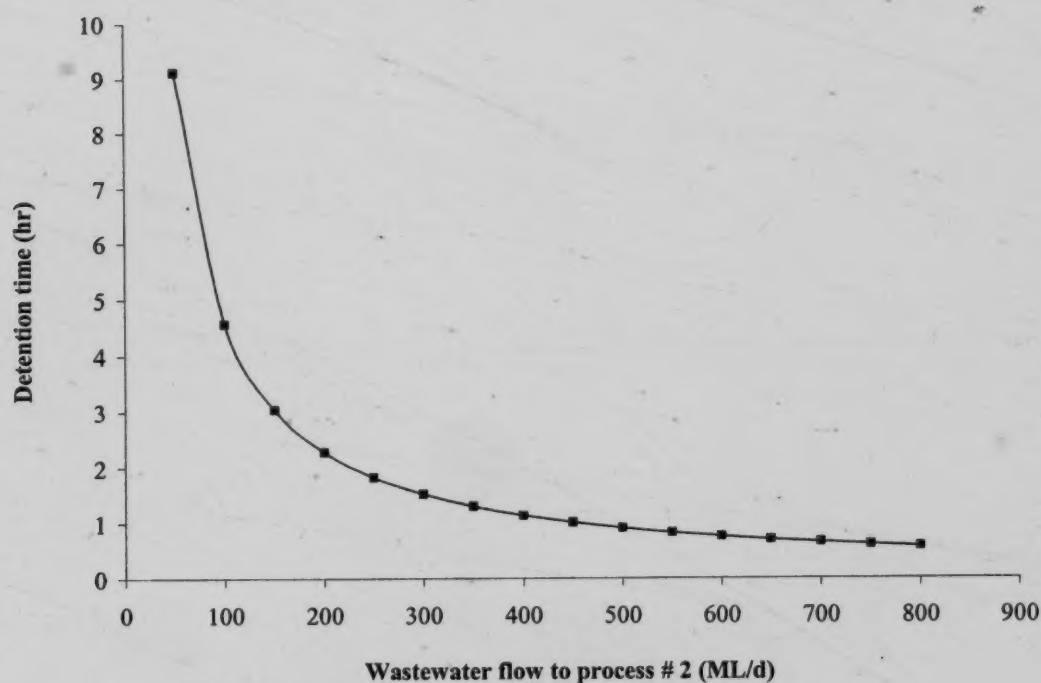
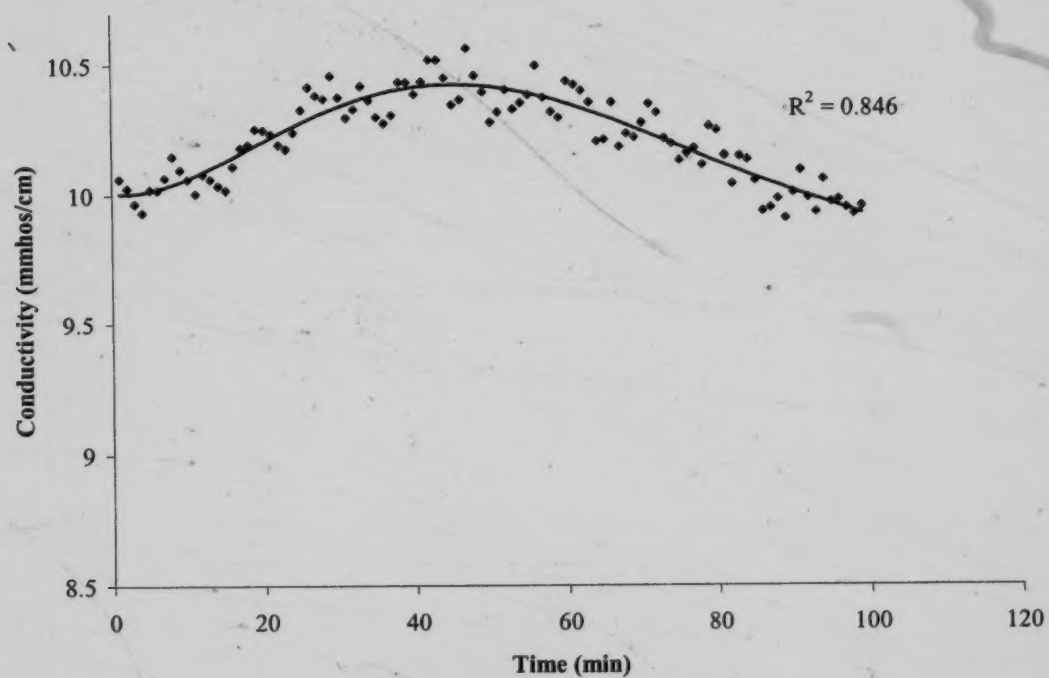


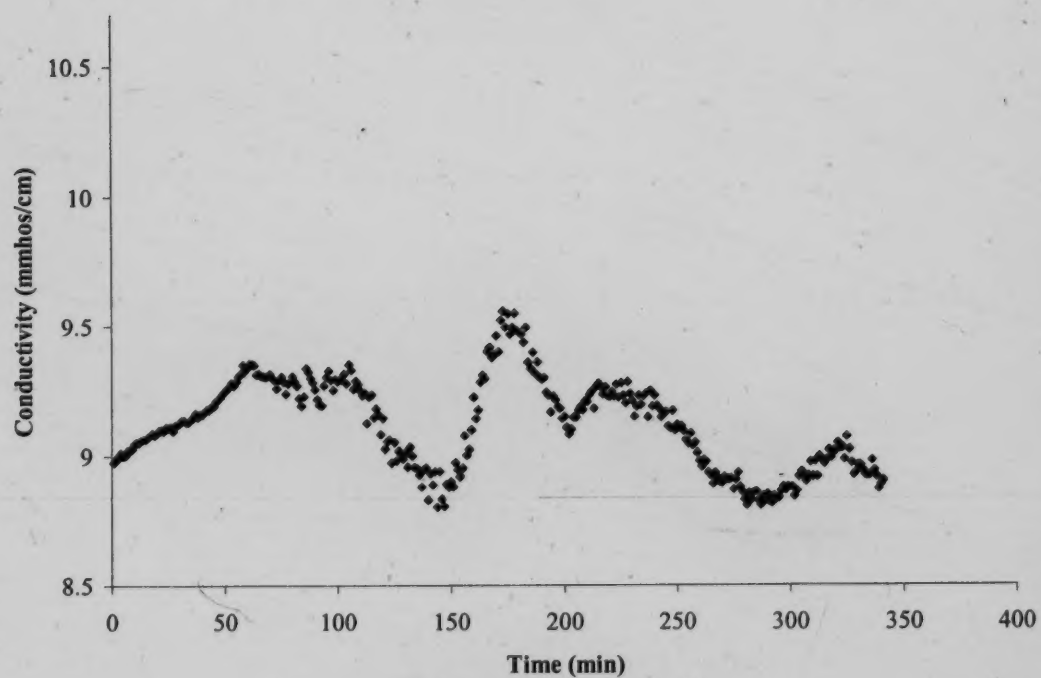
Figure 3.7 Theoretical detention time.

3.7.2 Tracer Studies

The theoretical residence time curve that is shown in Figure 3.7 never exists in practice, especially for full-scale sedimentation basins. The ideal settling plug-flow conditions are never attained in practice and all sedimentation tanks are subject to some degree of hydraulic turbulence, short circuiting, and density currents.



(a) First tracer study



(b) Second tracer study

Figure 3.8 Tracer studies.

Therefore, the actual detention time is likely to be less than the theoretical detention time calculated from the tank volume and the rate of flow. Two simple slug-input tracer studies for sedimentation tank #5 were conducted using water softening salt (brine) as a tracer. For both studies, a slug input of tracer was dumped at the influent channel and the conductivity of the effluent wastewater from the tank was measured at the same point that was used for sampling the effluent TSS and COD. Measurements of the conductivity were made every three seconds and were averaged over 60 seconds intervals as shown in Figure 3.8. The outcome of the first study is shown in Figure 3.8 (a) while Figure 3.8 (b) shows the outcome of the second tracer study. Time zero on the horizontal axis represents the time at which the slug was dumped.

For the first tracer study, the wastewater flow received by PST 2 was approximately 255 ML/d at the time of dumping the tracer. From the outcome shown in Figure 3.8 (a), it is evident that the peak concentration reached the effluent sampling point approximately 50 minutes (0.83 hr) from the time of dumping the tracer. The flow to PST 2 of 255 ML/d recorded at the time of dumping the slug for the first tracer study corresponds to a theoretical detention time of 1.8 hr.

For the second tracer study, the wastewater flow received by PST 2 was approximately 125 ML/d at the time of dumping the tracer. At the time of dumping the slug, the baseline conductivity was approximately 9 mmhos/cm. After approximately 60 minutes, a first peak of 9.35 mmhos/cm was observed. A second peak of 9.55 mmhos/cm appeared after 3 hours from the time of dumping the tracer. The flow of 125 ML/d at the time of dumping the tracer corresponds to a theoretical detention time of 3.65 hr.

It is clear from the outcomes of the tracer studies that the sedimentation tank suffers from short circuiting. The presence of short circuiting in full-scale sedimentation tanks has been observed in many studies in the literature. Tebbutt (1979) stated that the effects of hydraulic turbulence and density currents were observed to produce tracer peaks after less than one hour from full-scale sedimentation tanks in England with theoretical retentions of approximately 10 hours.

3.7.3 Neural Network Models Developed

The objective of the neural network modelling effort was to predict the TSS and COD of the effluent wastewater from the sedimentation tank. Two separate networks were trained for this task. One was trained to predict the effluent TSS and the other to predict the effluent COD. As suggested by Equation 3.1, the TSS network used the flow and influent TSS data as inputs. The COD network utilized the flow and influent COD data as inputs in order to predict the COD of the effluent wastewater from the tank.

When the characteristics of the effluent wastewater from the sedimentation tank are modelled, there should be a lag time between the effluent and the influent data used in modelling. This lag time is equal to the actual detention time of the tank and will depend on the flow volume entering the tank at that time. The outcomes of the tracer studies suggest that this lag time could be as low as less than one hour at high flows. Therefore, limited by the sampling interval of one hour that was used in the present study, the modelling was conducted using a lag (or delay) time that was equal to one hour.

In order to present the input data records to the neural network, a moving window of past records had to be used, however, the size of moving window had still to be determined. Using a very small window may hinder the ability of the network to provide adequate predictions because the whole picture is not fully seen by the neural network model. On the other hand, if a very large window size was to be used, the dimensionality of the input layer will increase which will necessitate the utilization of more parameters. In neural networks, this means more hidden nodes, and as a result, more connection weights. In the present study the size of the moving window was changed incrementally from 1-hour horizon to 5 hours with increments of one hour. This systematic approach, as will be shown later, allowed the selection of the time span that gives the best predictions and at the same time kept the model as parsimonious as possible.

3.7.4 Structure of the Neural Network Models

Much of the success in applying neural networks to engineering problems is mainly attributed to the achievements of the multi-layer feed-forward neural network architecture with back-propagation training algorithm (Lippmann, 1987). In this ANN architecture, information is processed in a forward manner through the network while the prediction error is propagated backwards through the network. A feed-forward neural network with back-propagation training algorithm was used for the modelling that was conducted in this study. Many researchers in the field of artificial neural networks suggest that it is usually unnecessary to use more than one hidden layer in a multi-layer feed-forward network and varying the number of hidden nodes in the one hidden layer is usually sufficient for delivering distinct results. In the present study only one hidden layer was used. A systematic approach was used to identify the size of the hidden layer. This approach is discussed in detail by Gamal El-Din and Smith (2002). In the present study, from the initial testing of the model, it was found that the smallest size of the hidden layer that allowed the network to converge was 3 nodes. The size of the hidden layer was increased incrementally from 3 nodes to 10 nodes with increments of one node and the parsimonious model that gave the best results was selected.

3.7.5 Training the Networks

In the modelling that was undertaken in this study, a batch-mode (weight updates were done after each epoch and not after each training pattern) back-propagation algorithm was used in the course of training and the network was saved at the point of minimum training error. The input data were scaled linearly into the range $<<0,1>>$ before being presented to the networks. From the initial model testing conducted in the present study using three different transfer functions by the hidden layer (logistic, tanh, and Gaussian), it was found that the logistic and tanh functions yielded the best results, and therefore, were used for further modelling efforts. A linear function in the output node was used as the initial testing of the model showed better results when using a linear function in the output node.

A systematic approach suggested by Gamal El-Din and Smith (2002) was used to determine the number of training epochs at which training was stopped. It will be seen later that this approach allowed the neural network to learn the important features from the data without memorizing it. In the present study training was stopped after 25, 50, 100, 150, 200, 300, 400, 500, 750, 1000, 1250, 1500, 1750, 2000, 2500, 3000, 3500, 4000, 4500, and 5000 epochs for every candidate model that was tested. Each time training was stopped, the model was tested against the validation data set. For the TSS modelling, 14 days of hourly data were available for modelling. The data for the period of July 3, 6:00 hours till July 5, 6:00 hours were used for validating the network and the rest of the data set was used for training. For the COD modelling, 10 days of hourly data were available for modelling, out of which, the two-days stretch of July 3, 6:00 hours till July 5, 6:00 hours was used as the validation data set and the rest of the data were used for training.

3.8 RESULTS OF THE NEURAL NETWORK MODELLING

The performance of the networks was measured using the coefficient of determination (R^2), which is a statistical indicator that compares the accuracy of the model to the accuracy of a trivial benchmark model wherein the prediction is just the mean of all the samples. A perfect fit would result in an R^2 value of "1", a very good fit near "1", and a very poor fit near "0". The coefficient of determination, R^2 , is mathematically described as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{i=n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{i=n} (y_i - \bar{y})^2} \quad [3.2]$$

where y_i is the actual output value, $\hat{y}_i$ is the output value predicted by the network, $\bar{y}$ is the mean of y values, and n is the total number of data records.

For each of the TSS and COD models, in order to try 5 different sizes of moving window, 8 different sizes of hidden layer and stopping training at 20 different numbers of epochs, a total of 800 ($5 \times 8 \times 20$) runs was conducted twice. Once when the logistic function was used by the hidden nodes and the second when the tanh function was used as the transfer function in the hidden layer. For each size of moving window, 8 different sizes of hidden layer were tried, and for each one of them, training was allowed for 20 different number of cycles.

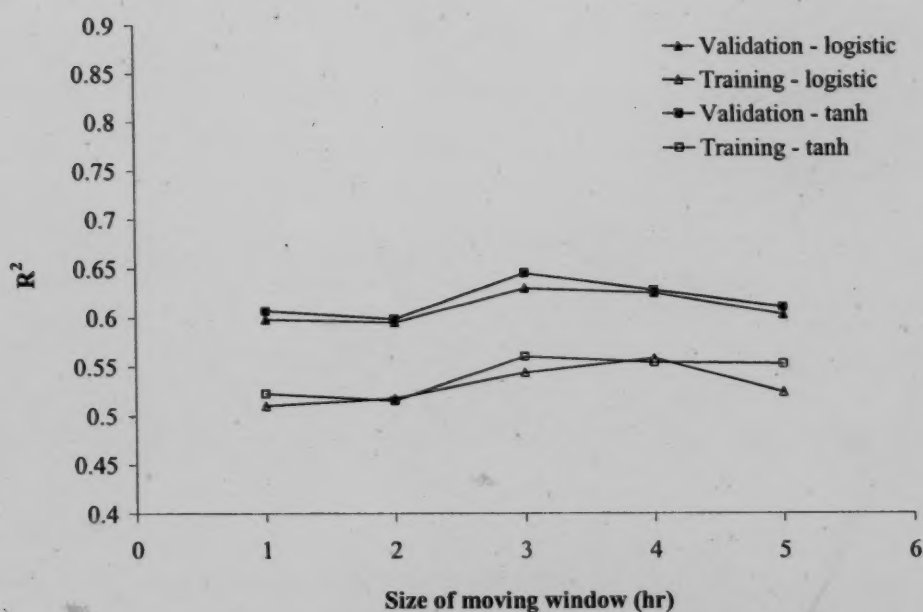


Figure 3.9 The effect of increasing the size of moving window.

It was found that for almost all the candidate models that were able to converge, as the number of epochs that was used in training increased, the value of the R^2 for the training set increased and the maximum value was obtained when training was allowed for 5000 epochs (the maximum number of epochs that was used). On the other hand, the maximum value of R^2 for the validation data set was always obtained at a much lower number of epochs (on one instance was as low as 50 epochs). Candidate models that were able to achieve a value of R^2 higher than 0.5 for both training and validation sets were

considered as being able to converge. For any candidate model that was able to converge, the model was allowed to train on the training data set for a number of epochs that allowed the model to generalize the best, i.e., the number of epochs after which the maximum R^2 value for the validation data set was obtained.

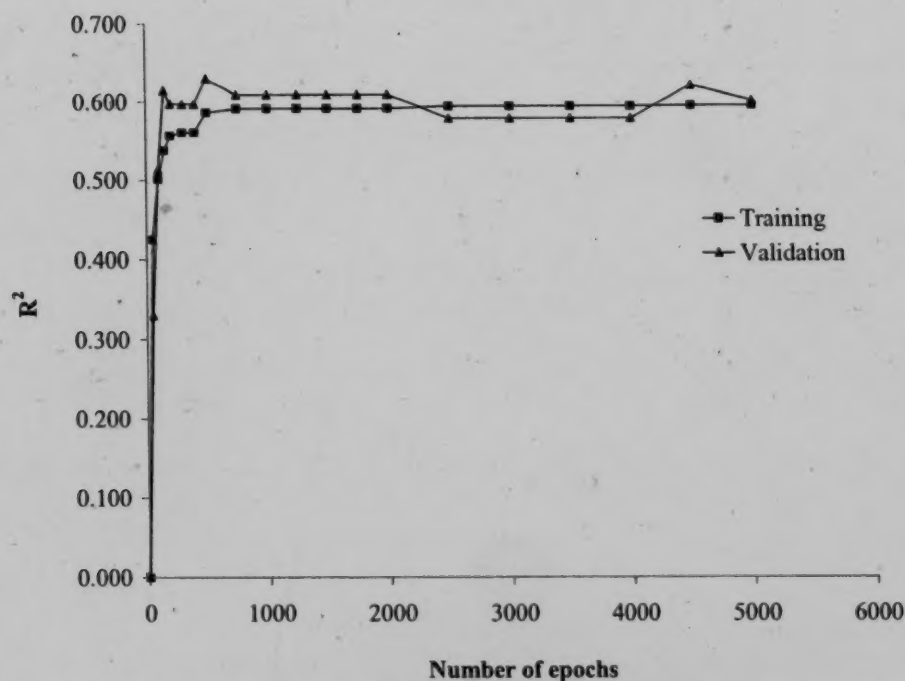


Figure 3.10 Results obtained by the TSS model.

When averaging the results of the candidate TSS models that were able to converge, the results are shown in Figure 3.9 from which it is evident that using a moving window of size 3 hours yielded the best results for the validation data set for the TSS model. Similar findings were obtained for the COD model. Among the candidate TSS models tested that used a 3 hour moving window, the model structure that used 9 hidden nodes with the tanh function as the transfer function yielded the best performance in terms of the ability of the network to learn the training data set and to generalize well when tested against the validation data set. The results obtained by this model structure are shown in Figure 3.10 from which it is clear that continuing training beyond 500 epochs did not improve the value of R^2 for validation and had a slight improvement in the value for the

training set. Therefore, it was decided to stop training of the final model for TSS after 500 epochs had elapsed and the model predictions for the effluent TSS are shown in Figure 3.11 and Figure 3.12 for the training and validation data patterns, respectively.

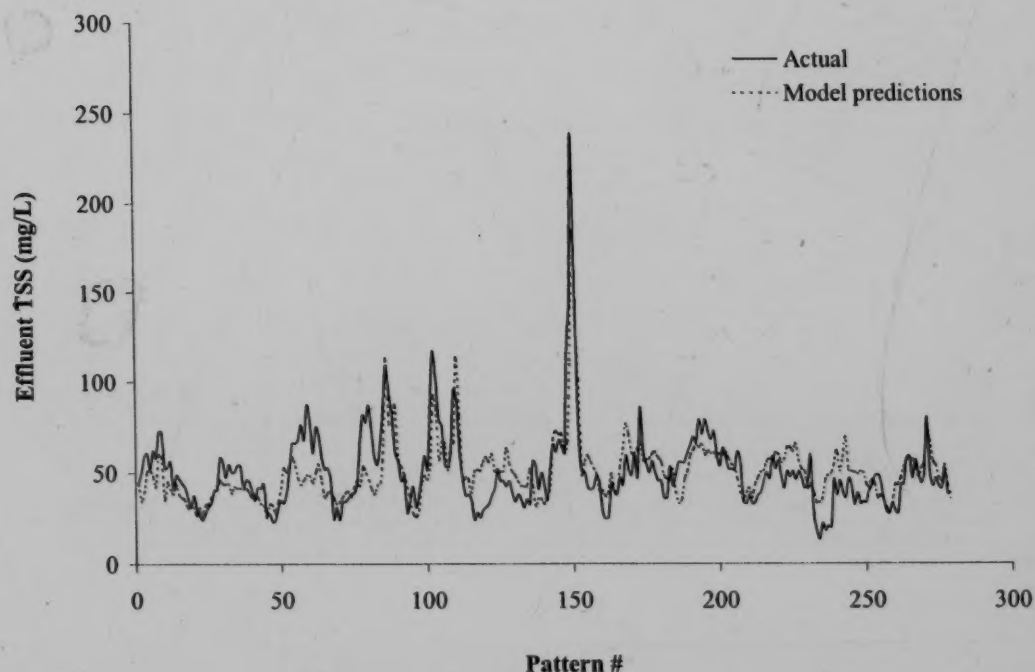


Figure 3.11 Predictions obtained by the TSS model for the training data set.

It is clear from Figure 3.12 that the neural network model for the TSS was able to generalize well when tested against the validation data set. Figure 3.11 shows that the actual effluent TSS values were generally underestimated by the model during dry weather periods of the first week of sampling, while were overestimated during dry weather periods of the second week. Both the flow and influent TSS during dry weather flow periods of the second week of sampling were higher than that of the first week, and despite that, the effluent TSS values were similar. The higher temperature encountered during the second week of sampling can explain the better removal of TSS. When the temperature of wastewater increase, the viscosity of the water will decrease, and as a result, the settling velocities of the particulates will increase and better removals will be attained.

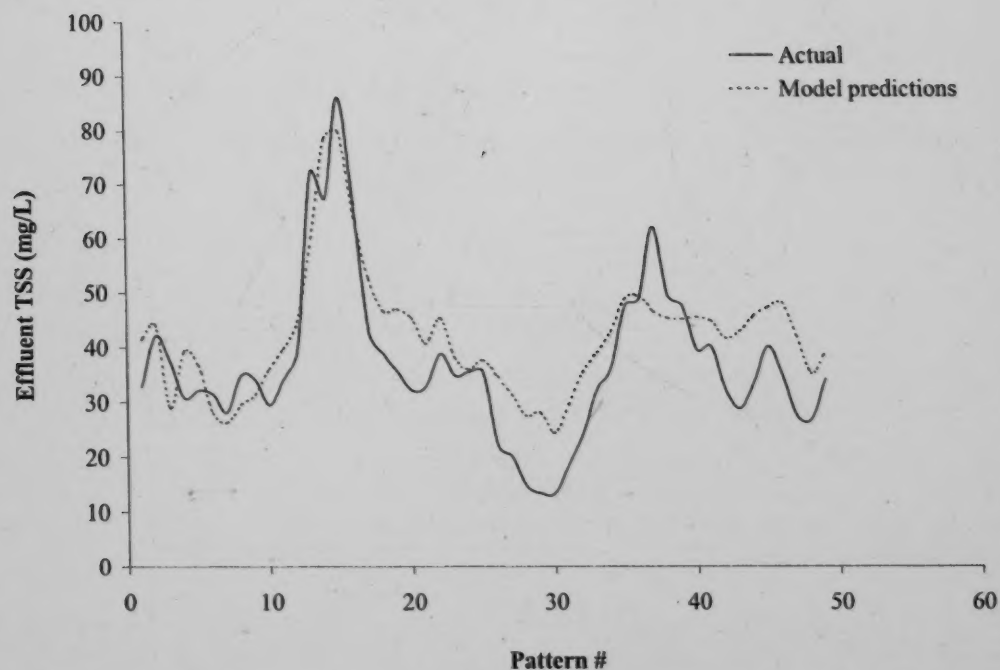


Figure 3.12 Predictions obtained by the TSS model for the validation data set.

The model predictions were better during wet weather conditions. The quality data of untreated wastewater always has an inherent random component which is more pronounced during dry weather conditions. When a rainfall event occurs, depending on the intensity and duration of the storm, the flow will increase and in most cases the influent TSS will also increase due to the first flush effect. During such events, the increase in influent TSS will mask the inherent random component within the data and as a result, it will be easier for the neural network to recognize the increase in both the flow and the influent TSS and predict the TSS of the effluent wastewater, and therefore, the prediction capability of the neural network model was much better during rainfall events. From a process control point of view, it is important to predict the performance of the tank during process upsets such as rainfall events.

Among the candidate COD models tested that used a 3 hour moving window, the model structure that used 9 hidden nodes with the tanh function as the transfer function

yielded the best performance. Figure 3.13 shows the results obtained by this model structure from which it is clear that allowing the network to train beyond 3500 epochs did not have any pronounced improvement in the ability of the model to generalize to the validation data set. Therefore, training of the final model for COD was stopped after 3500 epochs and the model predictions for the effluent COD are shown in Figure 3.14 and Figure 3.15 for the training and validation data patterns, respectively. When tested against the validation data set, the ability of the COD neural network model to generalize well is evident from Figure 3.15. The model was able to predict the decrease in effluent COD values during the course of rainfall events that were encountered during the sampling periods. The reason that the COD model gave better predictions than that of the TSS model is the strong correlation that existed between the influent and effluent COD data. The Pearson correlation coefficient was calculated between the effluent and influent COD data sets lagged by one, two, three, and four hours and was found to be 0.804, 0.779, 0.726, and 0.649, respectively.

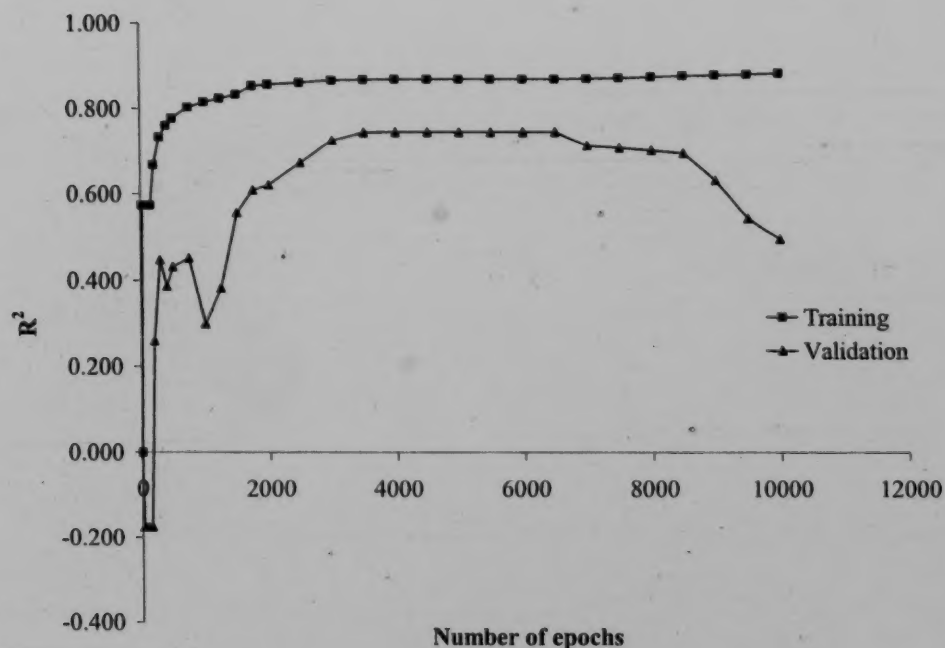


Figure 3.13 Results obtained by the COD model.

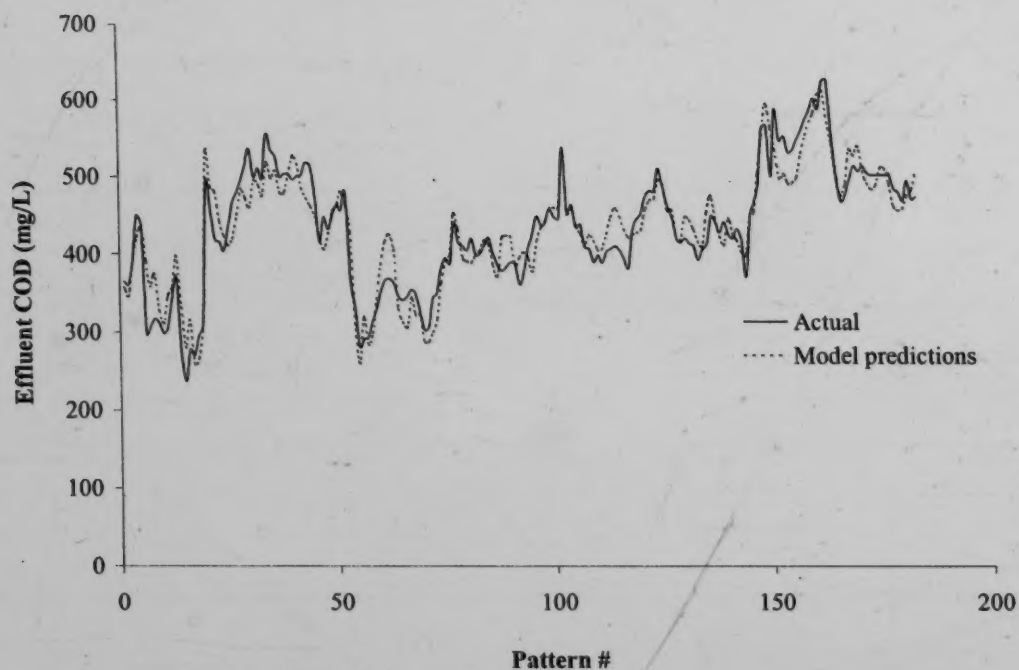


Figure 3.14 Predictions obtained by the COD model for the training data set.

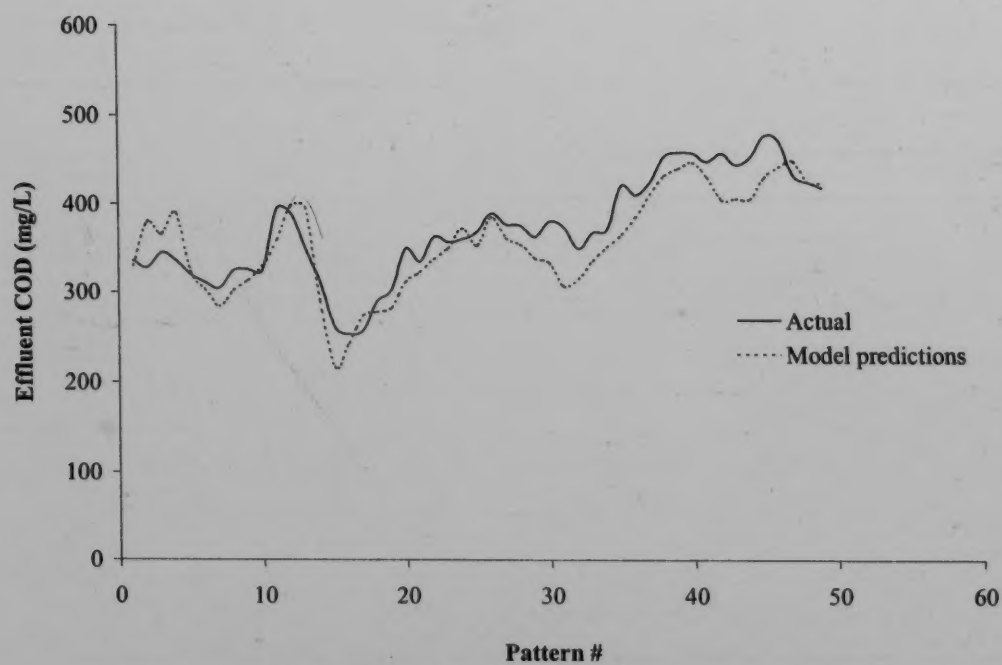


Figure 3.15 Predictions obtained by the COD model for the validation data set.

3.8.1 The Neural Network Models Against Traditional Regression Models

The prediction ability of the neural network models developed was tested against that of traditional regression models of the form of Equation 3.1. In order to do that, nonlinear regression was used to estimate the values of the parameters a , b , and c . The influent flow values recorded during the preceding three hours were averaged to calculate the surface overflow rate q . Similarly, the influent TSS and COD values measured during the preceding three hours were averaged to calculate the influent concentration S . The one-hour-ahead effluent TSS and COD values were used to calculate the fractional removal E . The R^2 value obtained by the regression model for the effluent TSS was 0.281, which is very low compared to the values obtained by the neural network models (0.586 for training and 0.629 for validation). The regression model for the effluent COD gave an R^2 value of 0.191, which is also much lower than that of the neural network model (0.868 for training and 0.744 for validation).

3.9 DISCUSSION

Models such as the ones presented here may have two potential uses. The first, for simulating the response of the sedimentation tank to different flow and influent quality profiles in order to choose between different operation strategies. The response of the sedimentation tank to different load conditions then can be used as an input to the models of the downstream biological processes. The second, for online control of the sedimentation basin. This could be of great value especially if chemicals are to be added to raw sewage to enhance TSS and COD reduction or for the purpose of phosphorus reduction. However, in order to use this current model for the purpose of online control, the inputs to the model have to be measured online.

The flow volume is measured online and can be incorporated into such a use. On the other hand, it is impossible to measure the TSS online. However, the quality of wastewater with respect to colloidal and suspended matter may be measured by turbidity. Linear relationships between total suspended solids and turbidity of wastewater treated

by an activated sludge system have been established. Although this type of relationship for untreated wastewater is much more difficult to establish, some researchers (Vanderborght and Wollast, 1990; Bertrand-Krajewski, 1992) have developed such relationships for the purpose of monitoring the total suspended solids on real-time basis.

Online COD analyzers are available, however, they are expensive and therefore more frequently used in the control of biological processes. They utilize ozone for the oxidation of the organic matter. Londong and Wachtl (1996) used a COD analyzer to monitor the influent untreated wastewater to a treatment plant in Germany. Instead of monitoring the COD itself, other parameters that can be correlated to the influent COD and in the same time can be measured online may be utilized. Among these parameters are the conductivity and turbidity. They have the ability to give information about the dissolved and suspended load in the wastewater flow. Hack and Kohne (1996) found very strong correlation between each of the conductivity and turbidity and the influent COD to a treatment plant. Finally, it should be mentioned that, although the models were trained on a limited amount of data, they still were able to give adequate predictions. If online analyzers were to be used this would facilitate the collection of more data that can be used to train and validate the model, and in such case, a scheduled retraining of the networks may be performed in order to check the validity of the models.

3.10 CONCLUSIONS

Most of the modelling of full-scale sedimentation tanks has been done using empirical models that utilize daily average data. In the present study a neural network model was developed to predict the dynamic behavior of a full-scale sedimentation tank. The model makes one hour in-advance predictions of the total suspended solids and chemical oxygen demand in the effluent wastewater from the sedimentation tank using flow data as well as influent TSS and COD data. A systematic approach was used in the building process of the model. The identified model for predicting the primary effluent TSS gave an R^2 value of 0.63 and 0.59 for the training and validation data sets, respectively. R^2 values of 0.87 and 0.74 were obtained by the model used to predict the

primary effluent COD for the training and validation data sets, respectively. The predictions of the effluent COD were better than that of the effluent TSS and the reason for that was the stronger correlation that was found to exist between the influent and effluent COD data. The few rainfall events that occurred during the sampling period had a pronounced effect on the quality of the effluent wastewater from the sedimentation tank. During these events, the total suspended solids for the effluent increased, however, the effluent chemical oxygen demand decreased. The models were able to distinguish those deviations from the dry weather values. Such models have the potential to be utilized for the purposes of simulation and online process control.

3.11 REFERENCES

1. Alarie, R.L., E.A. McBean, and G.J. Farquhar, "Simulation Modelling of Primary Clarifiers", J. Env. Eng., ASCE 106(EE2): 293-309 (1980).
2. APHA-AWWA-WEF, *Standard Methods for the Examination of Water and Wastewater*, 19th Ed. American Public Health Association, American Water Works Association, and Water Environment Federation (Washington, DC, USA: American Public Health Association, 1995).
3. Bertrand-Krajewski, J.L., "A Model for Solid Production and Transport for Small Urban Catchments: Preliminary Results", Wat. Sci. Tech. 25(8): 29-35 (1992).
4. Bertrand-Krajewski, J.L., M. Lefebvre, B. Lefai, and J. Audic, "Flow and Pollutant Measurements in a Combined Sewer System to Operate a Wastewater Treatment Plant and Its Storage Tank During Storm Events", Wat. Sci. Tech. 31(7): 1-12 (1995).
5. Djebbar, Y. and P.T. Kadota, "Estimating Sanitary Flows Using Neural Networks", *proc. Of IAWQ 19th Biennial International Conference*: Vancouver, Canada, 1998.
6. El-Baroudi, H.M., "Characterization of Settling Tanks by Eddy Diffusion", J. San. Eng., ASCE 95(SA3): 527-544 (1969).
7. Gamal El-Din, A. and Smith, D.W. "A Neural Network Model to Predict the Wastewater Inflow Incorporating Rainfall Events". Water Research 36(5): 1115-1126 (2002).
8. Hack, M. and M. Kohne, "Estimation of Wastewater Process Parameters Using Neural Networks", Wat. Sci. Tech. 33(1): 101-115 (1996).
9. Lippmann, R.P., "An Introduction to Computing with Neural Nets", IEEE ASSP Magazine. April: 4-22 (1987).
10. Metcalf & Eddy Inc., *Wastewater engineering treatment, disposal, and reuse*, 3rd ed. (NY, USA: McGraw-Hill, Inc., 1991).
11. Shiba, A.M. and Y. Inoue, "Dynamic Response of Settling Basins", J. Env. Eng., ASCE 101(EE5): 741-757 (1975).
12. Smith, R., "Preliminary Design of Wastewater Treatment Systems", J. San. Eng., ASCE 95(SA1): 117-145 (1969).

13. Steel, E.W., *Water supply and sewage*, 4th ed. (NY, USA: McGraw-Hill, Inc., 1960).
14. Takamatsu, T., M. Naito, S. Shiba, and Y. Veda, "Effects of Deposit Resuspension on Settling Basin", *J. Env. Eng., ASCE* 100(EE4): 883-903 (1974).
15. Tebbutt, T. and D. Christoulas, "Performance Relationships for Primary Sedimentation", *Wat. Res.* 9: 347-356 (1975).
16. Tebbutt, T.H.Y., "The Performance of Circular Sedimentation Tanks", *Water Poll. Control* 68: 467 (1969).
17. Tebbutt, T.H.Y., "Primary Sedimentation of Wastewater", *J. WPCF* 51(12): 2858-2867 (1979).
18. Thirumurthi, D., "A Break-Through in the Tracer Studies of Sedimentation Tanks", *J. WPCF* 41(11): R405-R418 (1969).
19. Vanderborght, J.P. and P. Wollast, "Continuous Monitoring of Wastewater Composition in Sewers and Stormwater Overflows", *Wat. Sci. Tech.* 22(10/11): 271-275 (1990).
20. Vondong, J. and P. Wachtl, "Six Years of Experience with the Operation of On-line Analyzers", *Wat. Sci. Tech.* 33(1): 159-164 (1996).
21. Voshel, D. and J. Sak, "The Effect of Primary Effluent Suspended Solids and BOD on Activated Sludge Production", *J. WPCF* 40(5): R203-R212 (1968).

CHAPTER 4. A COMBINED TRANSFER FUNCTION-NOISE MODEL TO PREDICT THE DYNAMIC BEHAVIOR OF A FULL-SCALE PRIMARY SEDIMENTATION TANK*

4.1 INTRODUCTION

In order to reduce the pollutional load on receiving streams, more stringent water quality standards will be applied in the near future, and therefore, many wastewater treatment plants will be forced to improve their performance in order to comply with these future standards. The conventional remedy to this problem is to enlarge the existing facility, which is costly and not always feasible. The alternative option is to improve the management and operation scheme of the plant.

Most of the existing treatment facilities have been designed using traditional time-invariant criteria that are derived from rather simple models that are identified by parameters obtained from steady state treatability studies and/or historical data (Novotny et al., 1992). Such facilities are then operated using an invariant (steady state) mode of operation which dictates that the input cannot exceed the bottleneck capacity of the treatment process and any excess is bypassed prior to the bottleneck and discharged to the receiving environment without treatment. In contrast, input into the system and the same treatment process dynamics are subject to high variability. The conflict between the modes of design and operation on one hand, and the modes and types of input and processes on the other, is one major reason why existing wastewater treatments plants often do not comply with applicable water quality standards (Novotny et al., 1992). Therefore, a conversion of operation to a dynamic real-time control (RTC) scheme may be a promising solution to this problem.

It is only recently that RTC systems have been used to control treatment plants (Capodaglio, 1994). The system requirements, objectives and components of RTC systems have been discussed elsewhere (Novotny et al., 1992; Capodaglio, 1994).

A part of this chapter has been accepted for publication. Gamal El-Din, A. and D.W. Smith, Water Research (February 2002).

Another part of this chapter has been accepted for publication. Gamal El-Din, A. and D.W. Smith, Journal of Environmental Engineering and Science (April 2002).

The ideal operational models in a RTC system for control of flow and/or pollution load discharges from urban sewerage and industrial wastewater treatment plants ought to be adaptive in response to both changes of the input waste loads and to the variation in the system parameters (Novotny et al., 1992).

One of the adaptive modelling technologies available to accomplish this task is the methodology developed by Box and Jenkins (Box and Jenkins, 1976), where both univariate and multivariate (transfer function) models may be used for analyzing time series data. These models are stochastic system models that are obtained by the system-identification strategy. They may retain the most important characteristics of the dynamic system they represent, without the need for extensive knowledge of the physical system being modelled. Since only observations of measurable inputs and outputs are needed for the identification of stochastic models, such models have to be developed specifically for the set of data under consideration, and then they constitute an adequate representation of the physical system only until a major change in the population generating the observations occurs (Capodaglio et al., 1990). When such changes do take place, stochastic models describing the system can be re-identified and/or re-estimated. This updating task is minimal in contrast to the tedious calibration process required by conventional deterministic models.

Stochastic models have been used in several applications to represent different types of dynamic systems with random features (Box and Jenkins, 1976). In the present paper, the Box and Jenkins methodology is utilized in order to study the performance of a full-scale primary sedimentation tank at the Gold Bar Wastewater Treatment Plant (GBWWTP), the largest treatment plant in the Edmonton area. The motivation behind the current modelling efforts is to improve the performance of the existing plant by exploring possible control strategies that might be implemented in the future. Firstly, it was necessary to study the stochastic nature of the influent and effluent streams, and the dynamic relationship between them.

4.2 STUDY

The Gold Bar Wastewater Treatment Plant (GBWWTP) was constructed in 1956 on the southwest shore of the North Saskatchewan River. The present capacity of the plant is 950 ML/d for primary treatment and 420 ML/d for secondary treatment. The plant treats domestic and industrial sewage from the City of Edmonton. There are two primary sections in the plant, Primary Settling Tanks Group 1 (PST 1), which includes settling tanks #1, 2, 3, and 4, and Primary Settling Tanks Group 2 (PST 2), which includes settling tanks #5, 6, 7, and 8. Primary sedimentation tank #5 was selected for sampling. Detailed description of the GBWWTP and the sampling that was conducted is provided elsewhere (Gamal El-Din and Smith, 2002).

Table 4.1 identifies the two data sets that were collected. The first survey was conducted between 6:00 a.m. June 28 and 6:00 a.m. July 5, 1999 and grab samples were taken manually every one hour. The laboratory work included total suspended solids (TSS) of primary influent and TSS of primary effluent. The second survey was conducted between 6:00 a.m. August 20 and 6:00 a.m. August 27, 1999, and as for data set #1, grab samples were collected manually every one hour. The survey program for the second week was expanded to include the following: (1) TSS of primary influent; (2) TSS of primary effluent; (3) chemical oxygen demand (COD) of primary influent; and (4) COD of primary effluent. For both surveys, the flow rate of the primary influent entering PST 2 was also recorded. All analyses were performed in triplicates in accordance with standard accepted practice (APHA, 1995). Figure 4.1 shows the data collected during the first survey. Data set #2 is shown in Figure 4.2 and Figure 4.3.

Table 4.1. Data obtained from God Bar Wastewater Treatment Plant.

Survey # (Data set #)	Sampling frequency	Dates	Data
1	Hourly	June 28-July 5, 1999	Flow rate, influent TSS, effluent TSS
2	Hourly	August 20-27, 1999	Flow rate, influent TSS, influent COD, effluent TSS, effluent COD

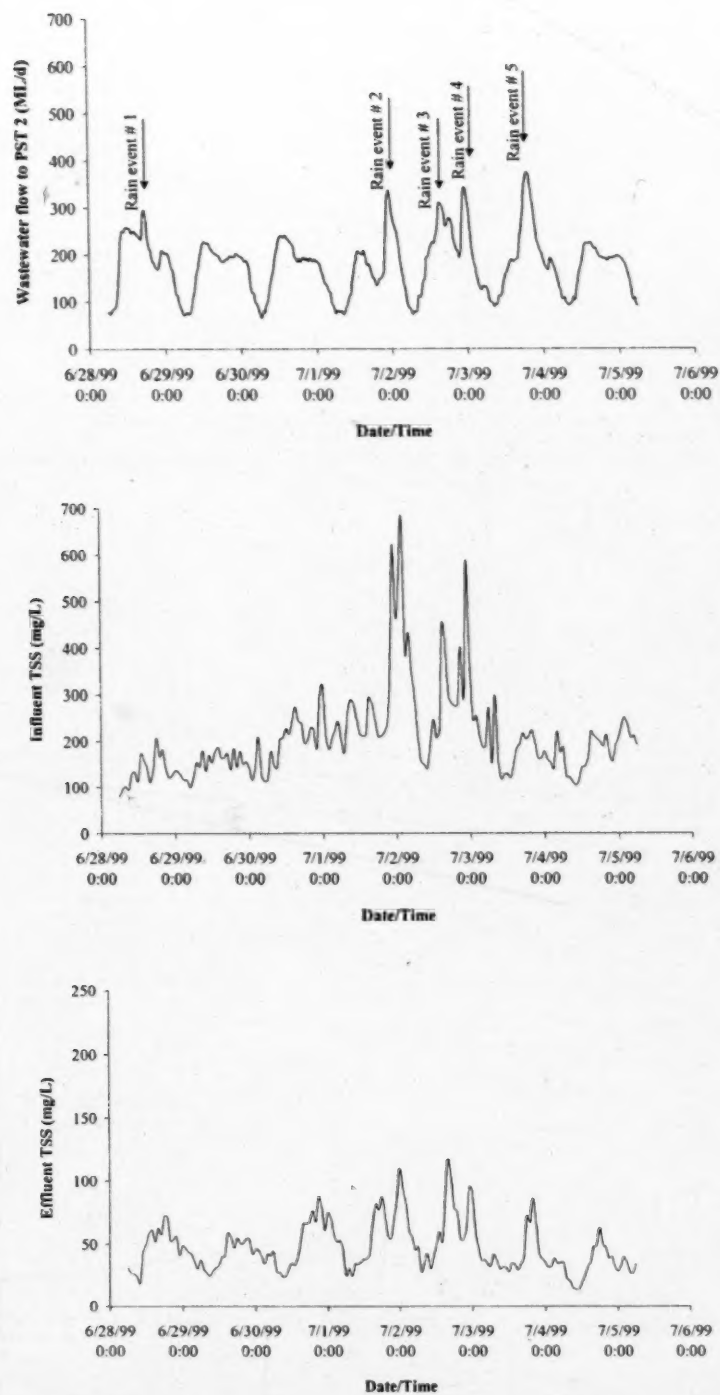


Figure 4.1 Influent and effluent hourly data for the 7-days survey of June-July 1999: data set #1.

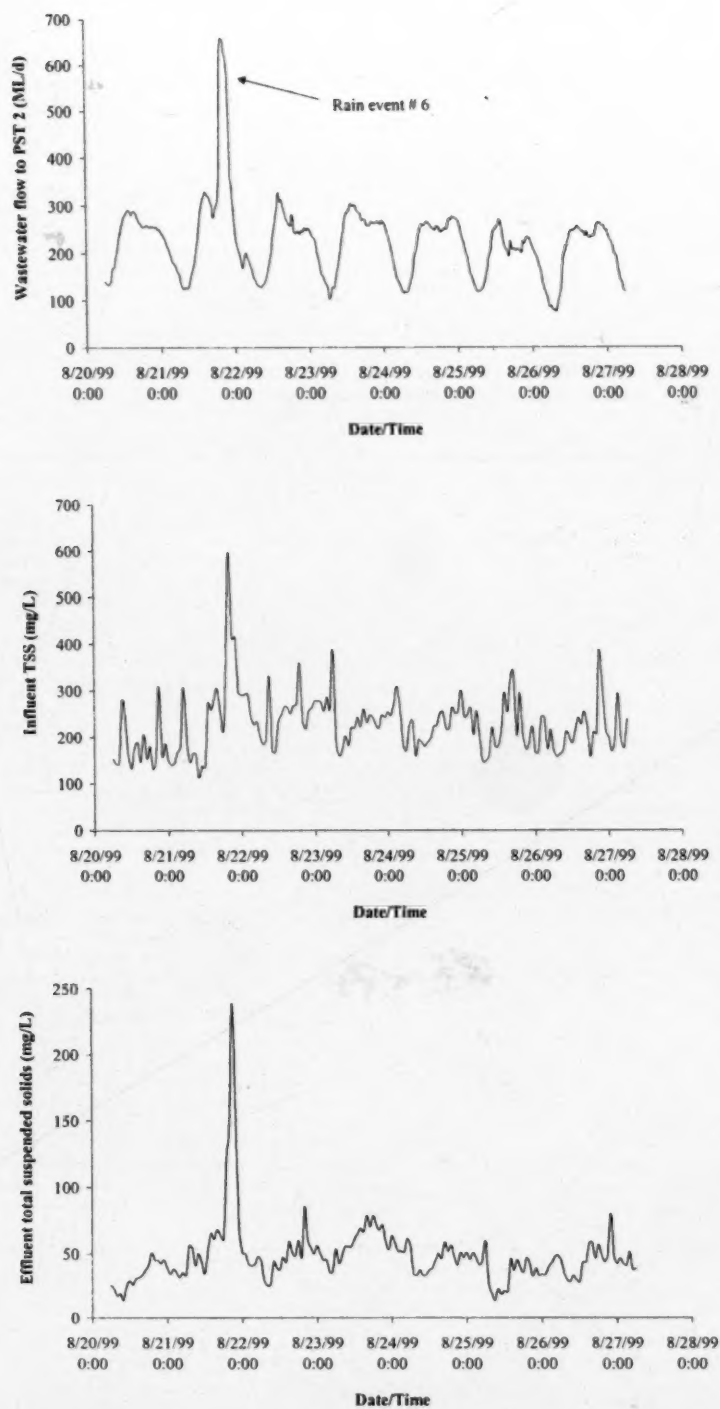


Figure 4.2 Influent hourly flow and TSS, and effluent hourly TSS data for the 7-days survey of August 1999: data set #2.

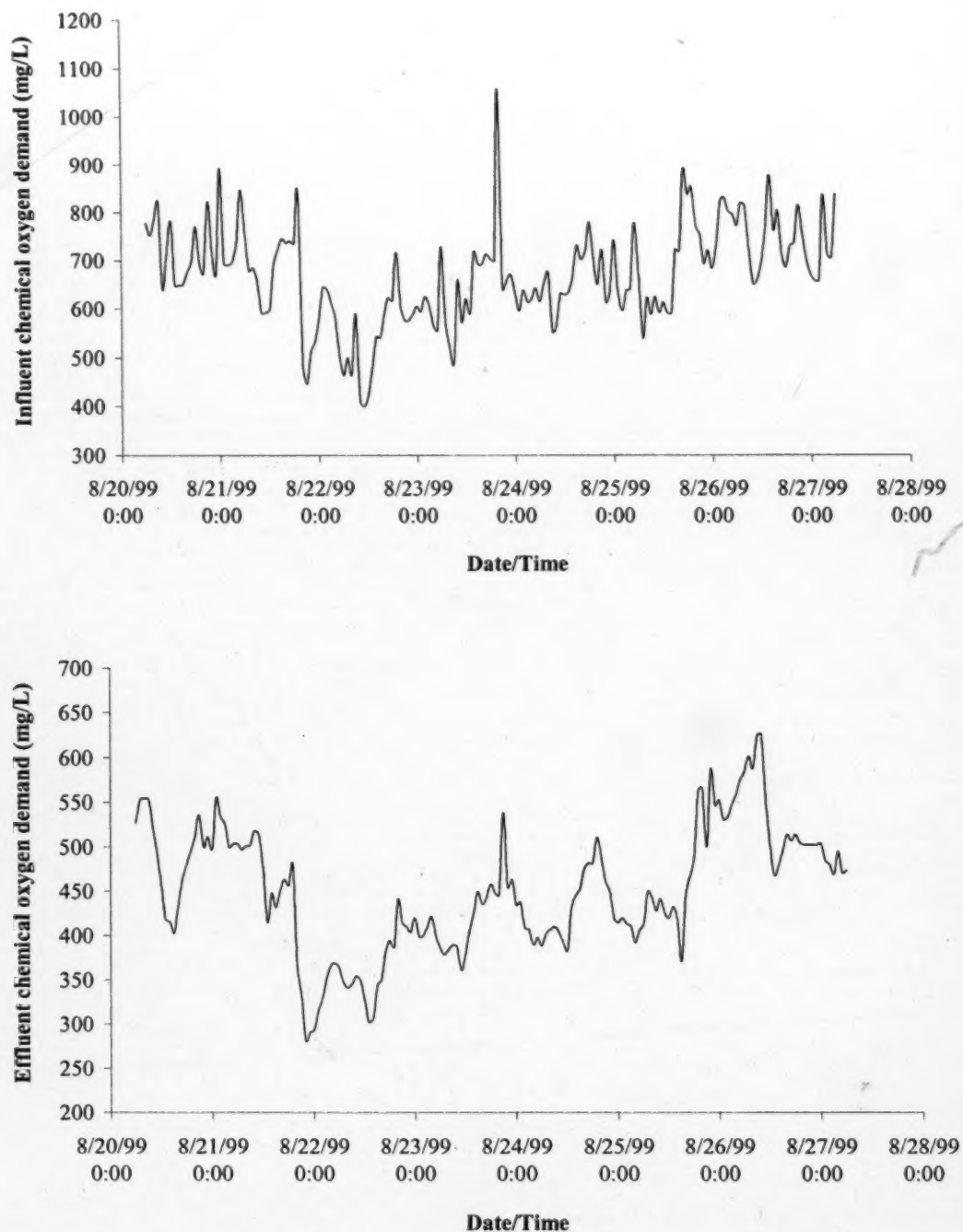


Figure 4.3 Influent and effluent hourly COD data for the 7-days survey of August 1999: data set #2.

4.3 QUALITATIVE DATA ANALYSIS

Some of the common patterns encountered in a time series are: an overall trend pattern (increase or decrease), a seasonal pattern, and a statistical pattern. The first two patterns are visual patterns that can generally be recognized when a time series is displayed in graphical form. On the other hand, statistical patterns cannot be identified by plotting the values of a time series and a statistical tool, like the Box-Jenkins method, has to be utilized. It is evident from Figure 4.1 to Figure 4.3 that there is no obvious trend in any of the series plotted. This was expected as the short duration over which each data set was collected (one week) did not allow any overall trend pattern to be obvious.

Seasonal behavior in a time series is simply the tendency of the series to repeat a certain pattern of behavior at regular time intervals called "seasons" (Box and Jenkins, 1976). The number of time series periods within a season is called the "periods per season". In the current study, seasonal behavior is expected due to the strong diurnal variation in wastewater flow data, and because hourly data were collected, the number of periods per season (denoted "S" in time series literature) is 24. One would expect relationships to occur between observations for successive hours in a particular day and between the observations for the same hour in successive days. Therefore, the situation is somewhat like that in a two-way analysis of variance model. Figure 4.1 to Figure 4.3 show an apparent seasonal pattern in the flow series, however, a much less evident seasonal pattern exist in the quality series (TSS and COD).

During the first survey, five rainfall events occurred and are indicated in Figure 4.1. Only one rainfall event occurred during the second week of sampling and is indicated in Figure 4.2. In the present study, the influent TSS increased substantially during most of the six rainfall events that were encountered during the sampling periods. During rainfall event # 6, which occurred during the second survey, the influent COD values had a steep drop from approximately 730 mg/L to around 450 mg/L, then increased gradually after the end of the event.

4.4 MODEL DEVELOPMENT

In the following sections a brief description of the model development is provided. Detailed conceptual and mathematical representation of the Box-Jenkins methodology can be found elsewhere (Box and Jenkins, 1976).

4.4.1 The Box-Jenkins Methodology

When representing the behaviour of a time series by the Box-Jenkins methodology, two general approaches may be used: the linear filter model approach and the transfer function model approach. The linear filter model approach is based on the idea that a time series in which successive values are highly dependent can be usefully regarded as generated from a series of uncorrelated independent "shocks" a_t , which are random drawings from a fixed distribution, usually assumed normal and having mean zero and variance σ_a^2 . Such a sequence of random variables $a_t, a_{t-1}, a_{t-2}, \dots$ is called a "white noise process".

A "linear filter" is a model that transform the white noise process a_t to the process that generated the time series, z_t , and can be represented mathematically by the equation $z_t = \psi(B)a_t$. This transformation is accomplished through the operator

$$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots = \sum_{j=0}^{\infty} \psi_j B^j \quad \text{with } \psi_0 = 1 \quad [4.1]$$

where B is the backshift operator such that $B^j a_t = a_{t-j}$. In order to have a parsimonious representation of the stochastic process represented by Equation 4.1, it is usually advantageous to write

$$\psi(B) = \frac{\theta(B)}{\phi(B)} \quad [4.2]$$

where $\theta(B)$ is the moving average operator of the stochastic model, and is defined as $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$; $\phi(B)$ is the auto-regressive operator of the stochastic model, and is defined as $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$; p and q are the orders of the stochastic model.

The linear filter model can represent any univariate Auto-Regressive Integrated Moving-Average (ARIMA) (p, d, q) model where d is the order of regular differencing needed to achieve stationarity. ARIMA models for time series with regular seasonal fluctuations have the general notations $ARIMA(p, d, q) \times (P, D, Q)_S$. The term (p, d, q) gives the order of the nonseasonal part. The order of the seasonal part is given by the term $(P, D, Q)_S$ where S is the number of observations in a season (24 in the current study). For example, the notation $ARIMA(1, 0, 2) \times (1, 0, 1)_{24}$ describes a seasonal ARIMA model for hourly data with the following mathematical form:

$$Y_t = \mu + \frac{(1 - \theta_1 B - \theta_2 B^2)(1 - \Theta_1 B^{24})}{(1 - \phi_1 B)(1 - \Phi_1 B^{24})} a_t \quad [4.3]$$

where μ is a mean term; $\theta_1, \theta_2, \phi_1$ are the regular moving average and auto-regressive parameters; Θ_1, Φ_1 are the seasonal moving average and auto-regressive parameters.

In contrast to ARIMA models, which describe the behaviour of single time series in terms of white noise, transfer function models can represent more complex systems in which the output is the stochastic response to one or more measured input series. The general form of a transfer-function noise model for the single input case is

$$Y_t = v(B)X_{t-b} + N_t \quad [4.4]$$

where Y_t is the output series at time t ; $v(B)$ is defined as $v(B) = (v_0 + v_1B + v_2B^2 + \dots)$ and is known as the impulse response function (it is the transfer function part of the model); X_{t-b} is the input series at time $t-b$, where b is a delay parameter; and N_t is a noise process at time t , defined by the linear filter $N_t = \psi(B)a_t$ and known as the stochastic model component (it is the noise part of the overall model). For the multiple input case, the model is

$$Y_t = v_1(B)X_{1,t-b_1} + v_2(B)X_{2,t-b_2} + \dots + N_t \quad [4.5]$$

The transfer function can be written

$$v(B) = \frac{\omega(B)}{\delta(B)} \quad [4.6]$$

where the numerator $\omega(B) = (\omega_0 - \omega_1B - \dots - \omega_sB^s)$; the denominator $\delta(B) = (1 - \delta_1B - \dots - \delta_rB^r)$; s and r are the orders of the polynomials. Combining Equations 4.5 and 4.6 yields

$$Y_t = \frac{\omega_1(B)}{\delta_1(B)}X_{1,t-b_1} + \frac{\omega_2(B)}{\delta_2(B)}X_{2,t-b_2} + \dots + N_t \quad [4.7]$$

Models represented by Equation 4.7 are usually called "transfer-function noise" models. The general approach for building such models of this type consists of identification, estimation (or fitting), and diagnostic checking.

4.4.2 Model Identification and Estimation

In the current study, the identification of a combined transfer function-noise model employed two separate steps; the first was the identification of the transfer function part of the model and the second was the identification of the stochastic part of the model. The first step was accomplished by calculating the sample cross correlation function $r_{xy}(k)$ at various lags k (see for example Figure 4.4), and then comparing it to theoretical impulse response functions of different orders in order to obtain some idea of the delay parameter b and the orders r and s of the operators in the transfer function (Box and Jenkins, 1976).

Before a cross correlation function between an output and an input series was calculated, both series were transformed using the same linear filter that produces a white noise having the input series as its input. This transformation process is called "pre-whitening" and was first introduced in 1976 by Box and Jenkins (1976). In cases when the estimated impulse response function suggested the consideration of more than one model, candidate models were estimated and diagnostic checking was performed on them in order to select the best model representing the system.

In the modelling effort presented in this paper, models were estimated using the maximum likelihood method outlined by Box & Jenkins (1976), in which the likelihood function is maximized via nonlinear least squares iterations. After a satisfactory model for the transfer function part has been identified and estimated, study of the sample autocorrelation and partial-autocorrelation functions of the residuals N_t in Equation 4.7 was used to identify the ARIMA model that represented the noise part at the output (see Box & Jenkins (1976) for details).

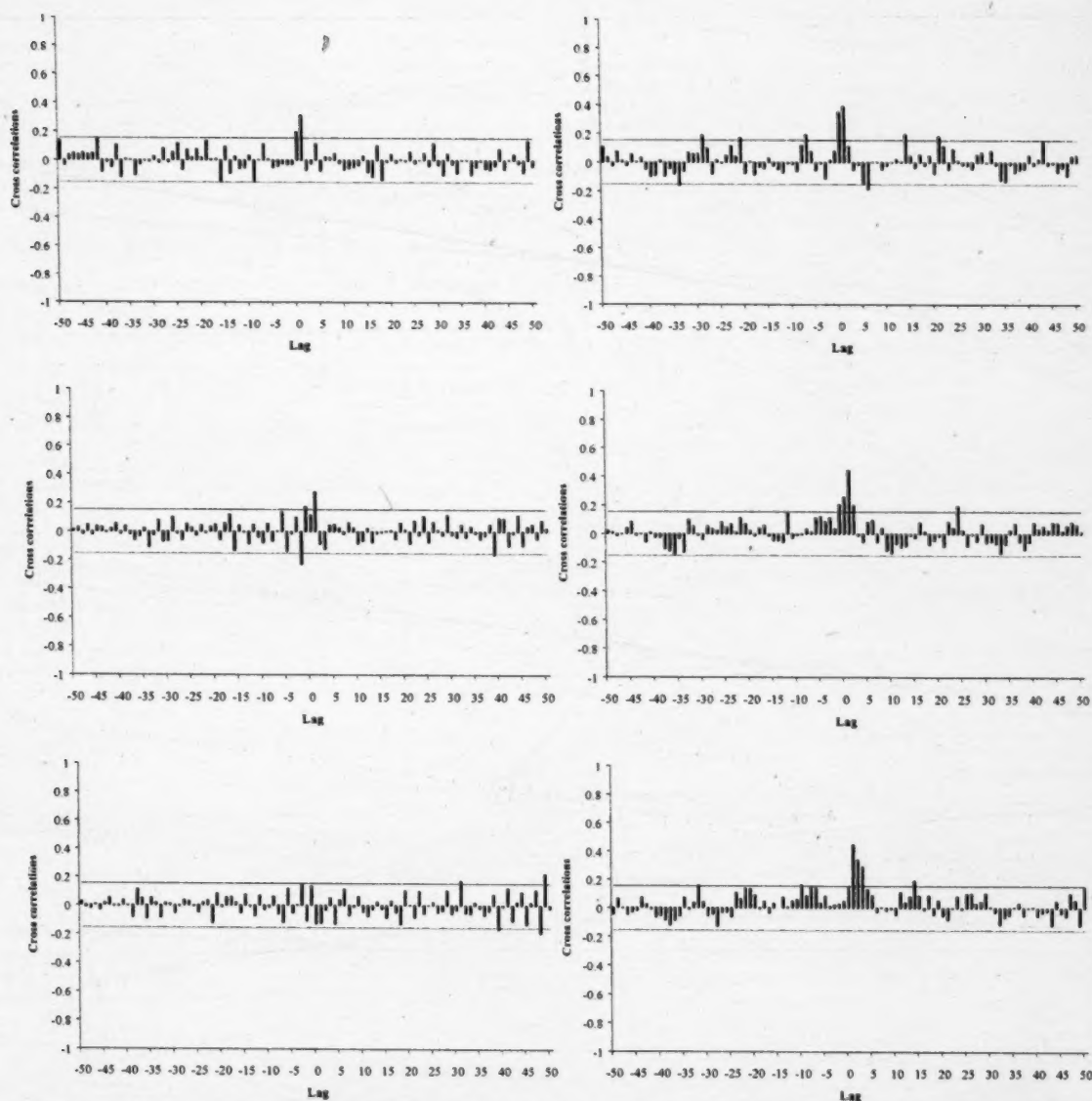


Figure 4.4 Cross correlation functions for prewhitened variables. Top left graph is flow and effluent TSS - survey #1; top right graph is influent TSS and effluent TSS - survey #1; middle left graph is flow and effluent TSS - survey #2; middle right graph is influent TSS and effluent TSS - survey #2; bottom left graph is flow and effluent COD - survey #2; bottom right graph is influent COD and effluent COD - survey #2. Solid lines represent the 95% confidence limits of two standard deviations.

4.4.3 Diagnostic Checking

After a model had been identified and estimated, it was checked to see whether it was an adequate model for the series. This step is called "diagnostic checking" of the model. Diagnostics that are applied to the fitted model include residual diagnostics and parameter diagnostics. In the current study, a variety of checks were applied to each model, and the test results were considered as a group. One technique which can be used for diagnostic checking is "overfitting". After the identification of what is believed to be the correct model, a more elaborate model, that contains additional parameters covering feared directions of discrepancy, is fitted to the data in order to put the identified model in jeopardy. In the present study, when a model was overfit, only one parameter was overfit at a time; numerator and denominator parameters were not overfit simultaneously.

4.4.3.1 Parameter Diagnostics

Parameter diagnostics included parameter confidence limits and correlations between parameters. In the current study, the 95% confidence limits (two standard errors) of a parameter were used to test the importance of including this parameter in the model. If the 95% confidence range included zero, then there is a strong possibility that the true value of the parameter is in fact zero (i.e., the parameter is not significant). A relatively high correlation between two parameters may indicate that one of them may probably be eliminated without affecting the adequacy of the model, and therefore, examining the measure of correlation between parameters was helpful in determining if a model was overspecified.

4.4.3.2 Residual Diagnostics

The statistical assumptions about the random error component a_t , implied by the theoretical Box-Jenkins methodology are such that the model residuals should be white noise, in other words, should be uncorrelated and normally distributed around a zero mean. Residual diagnostics are tools by which these assumptions can be tested. Further

more, models that have met these assumptions are compared using closeness-of-fit statistics applied to the residuals. Some of the statistics that can be computed as part of the residual diagnostics are the residual mean (mean error) and mean percent error.

Assuming that the form of the model is correct, the estimated autocorrelations of the residuals would be uncorrelated and distributed approximately normally about zero (Box and Jenkins, 1976). Therefore, correlograms of the residuals (see Figure 4.5 for an example) were examined for correlations greater than two standard deviations since large correlations may have indicated model inadequacies, especially if they were at lower lags.

Because of the fact that individual autocorrelations may fall within acceptable limits, but, for example, the first 20 autocorrelations combined as a group may be too high, a white noise check that considers groups of residual autocorrelations was important. In order to test the null hypothesis that a current set of autocorrelations is white noise, test statistics were calculated for different total numbers of successive lagged autocorrelations using the Ljung-Box formula (Ljung and Box, 1978)

$$Q = n(n+2) \sum_{k=1}^m \frac{r_k^2}{(n-k)} \quad [4.8]$$

where m is the total number of lagged autocorrelations under investigation and r_k is the sample autocorrelation of the residuals at lag k (Box and Jenkins, 1976). The test is made by comparing the Q -statistic with a critical test value (the chi-square value) and if the Q -statistic is larger than the critical test value, then one would conclude with a certain degree of confidence that the residual autocorrelations, being tested as a whole, are significant. The Q -statistic is compared to the chi-square value at $(m-P)$ degrees of freedom, where P is the number of parameters estimated. The Q -statistic was calculated for $m = \{12, 24, 36, \text{ and } 48\}$.

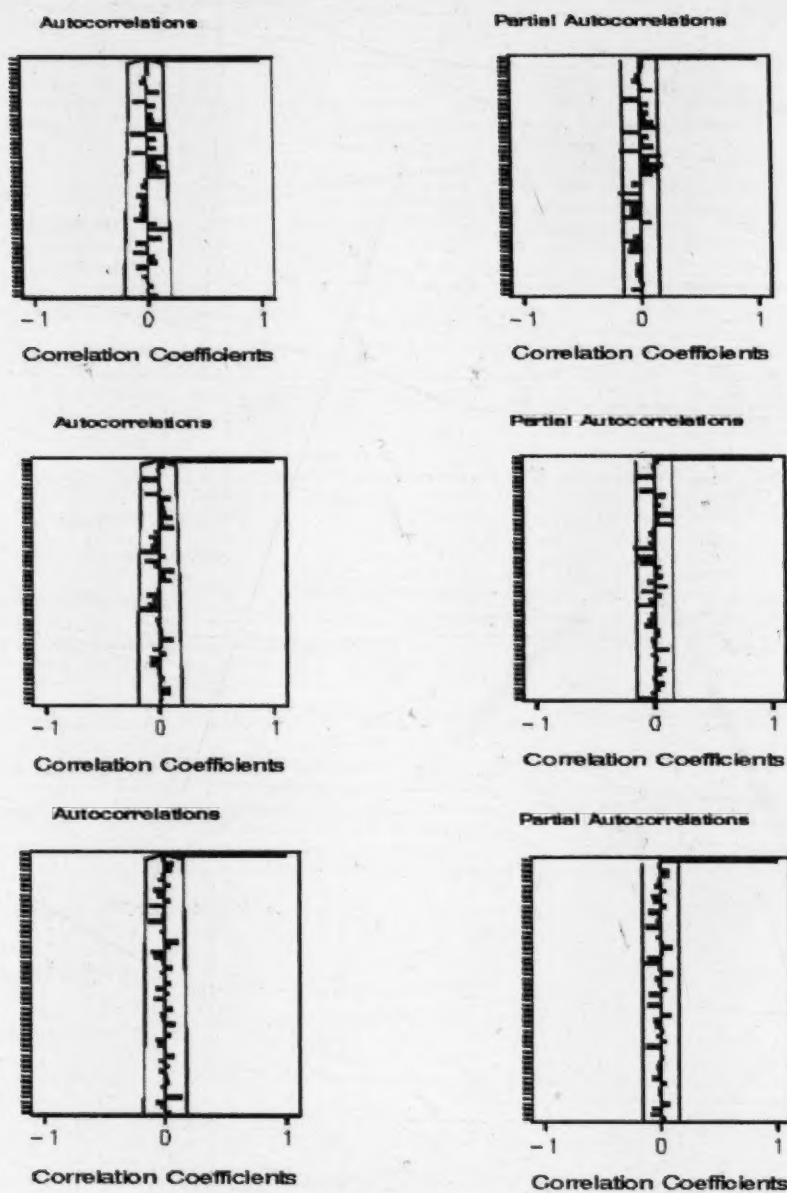


Figure 4.5 Autocorrelation and partial autocorrelation functions for residuals. Top graph is from model M-1; middle graph is from model M-2; bottom graph is from model M-3. Lags from 0 to 50 are shown on the vertical axes. Solid lines represent the 95% confidence limits of two standard deviations. For description of the models, see Table 4.2.

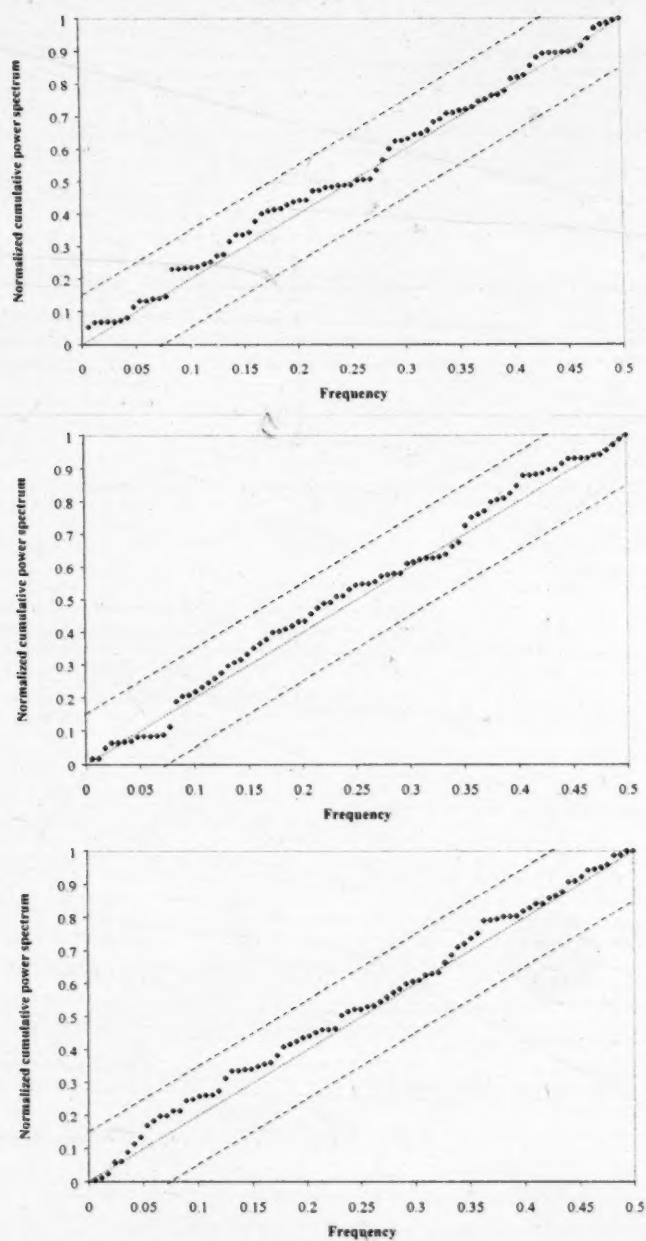


Figure 4.6 Cumulative periodogram check on residuals. Top graph is from model M-1; middle graph is from model M-2; bottom graph is from model M-3. Lags from 0 to 50 are shown. Dashed lines represent the 95% confidence limit lines of the Kolmogorov-Smirnov white noise test. For description of the models, see Table 4.2.

When modelling seasonal time series like the ones encountered in the present study, it may be feared that the periodic characteristics of the series have not been adequately taken into account. Such departures from randomness most probably will not be identified by the correlogram of the residuals because periodic effects will typically dilute themselves among several autocorrelations (Box and Jenkins, 1976). On the other hand, the periodogram is a device that is especially designed for the detection of periodic patterns in a background of white noise. It is another way of analyzing a time series based on the assumption that it is made up of sine and cosine waves with different frequencies (Box and Jenkins, 1976). This device is used by the Box-Jenkins methodology to provide an additional residual check that is strongly recommended when dealing with seasonal series, and hence, was one of the checks that were used in the current study. The definition of the periodogram assumes that the frequencies are harmonics of the fundamental frequency $1/n$ where n is the number of residuals. If this assumption is relaxed and the frequency is allowed to vary continuously in the range 0–0.5 cycles, the periodogram is then referred to as the sample power spectrum. It has been shown by Bartlett (1955) that the power spectrum for white noise has a constant value $2\sigma_a^2$ over the frequency domain 0–0.5 cycles where σ_a^2 is the variance of the white noise. Therefore, for a theoretical white noise process, if the normalized (with respect to σ_a^2) cumulative power spectrum is plotted against the frequency f , one would have a theoretical straight line running from (0, 0) to (0.5, 1). If the model is adequate, then the plot of the estimated normalized power spectrum against the frequency f (see Figure 4.6 for an example) would be scattered about the theoretical straight line joining the points (0, 0) and (0.5, 1). Using the Kolmogorov-Smirnov white noise test, 95% confidence limit lines were placed about the theoretical line (Box and Jenkins, 1976) (see Figure 4.6 for an example).

The normality of residuals was checked by examination of the histogram (see Figure 4.7 for an example) and normal probability plot (see Figure 4.8 for an example) of the residuals.

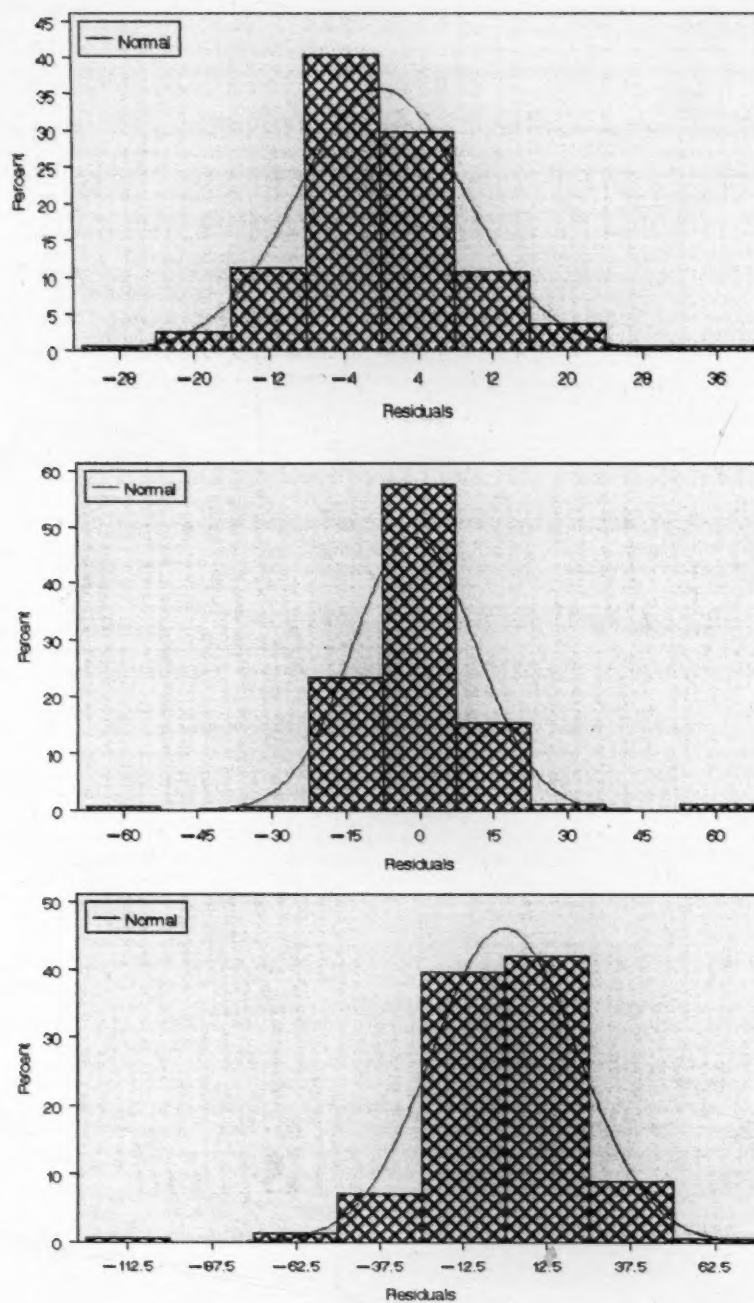


Figure 4.7 Histogram of the residuals. Top graph is from model M-1; middle graph is from model M-2; bottom graph is from model M-3. For description of the models, see Table 4.2.

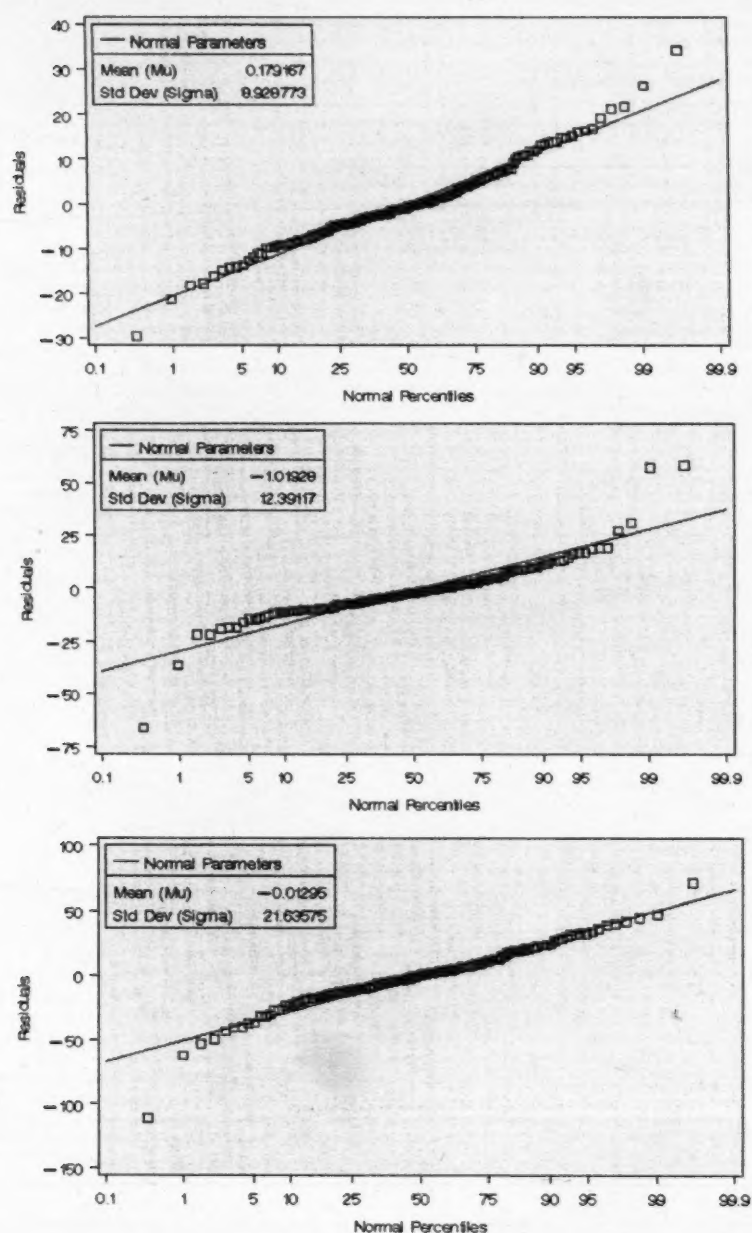


Figure 4.8 Normal probability plot of the residuals. Top graph is from model M-1; middle graph is from model M-2; bottom graph is from model M-3. For description of the models, see Table 4.2.

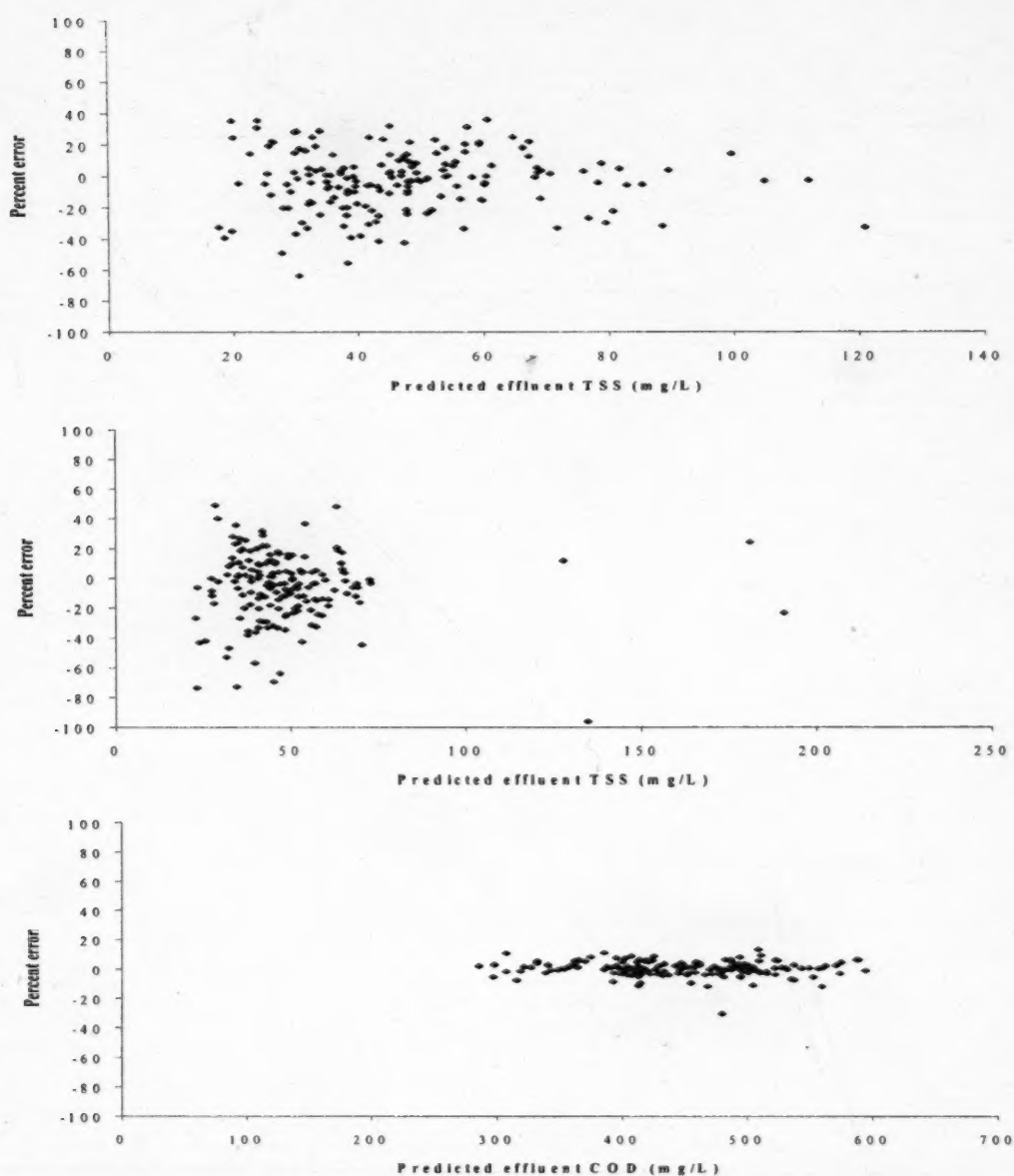


Figure 4.9 Residuals vs. predicted values. Top graph is from model M-1; middle graph is from model M-2; bottom graph is from model M-3. For description of the models, see Table 4.2.

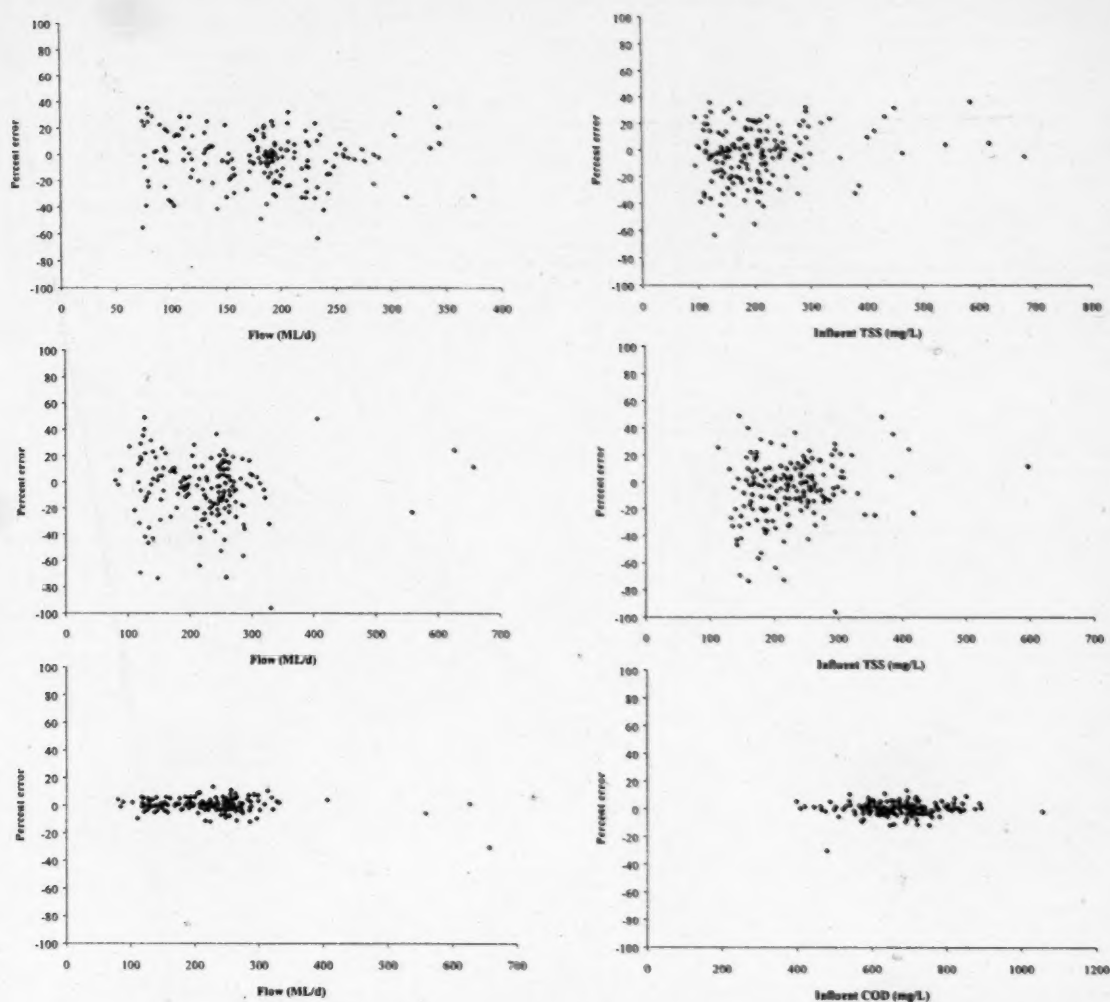


Figure 4.10 Residuals vs. input series. Top graphs are from model M-1; middle graphs are from model M-2; bottom graphs are from model M-3. Left graphs are residuals vs. flow; right graphs are residual vs. influent TSS (influent COD in the case of model M-3). For description of the models, see Table 4.2.

The residuals were also checked for homoscedasticity (constant error variance over all observations). This was done by examining a plot of the residuals versus the fitted values (see Figure 4.9 for an example). Finally, the independence of the residuals from

the input series was determined by examining a plot of the residuals versus the input series (see Figure 4.10 for an example) for any evidence of trends.

4.4.3.3 Closeness-of-Fit Statistics

Among the closeness-of-fit statistics are the mean absolute error, residual standard error, mean absolute percent error, and the index of determination " R^2 ". These are descriptive statistics that are useful for comparing different models that all passed the diagnostic checking step. For each candidate model that has been tested, all of the above mentioned statistics were calculated. In addition, plots of the correlogram, periodogram, histogram, and normal probability of residuals were drawn and white noise checks of the residuals were conducted in order to check the validity of the models.

4.5 QUANTITATIVE DATA ANALYSIS

It is the goal of this section to describe the building of a useful stochastic dynamic model which explains how and to what extent influent flow rate, TSS and COD and noise affect effluent TSS and COD. Before turning to transfer function models, one should see how much of the variation in Y_t can be explained by a stochastic time series model alone, which does not rely on any input variables as a predictor and it would be disappointing if a combined transfer function-noise model cannot do better (Box and Jenkins, 1976). Later in this section, it will be seen that the addition of a transfer function component will improve the prediction, and with the use of a transfer function model by itself (no noise component), the performance was worse than with a noise model by itself (no transfer function component).

In all the modelling that has been conducted, time series data were split into two parts, one for estimating the model parameters (i.e. calibrating the model) and the other for validating (i.e. verifying) the model. After a model has been estimated, the validation data set was used to judge the accuracy of the forecasts generated by the estimated model. This was done by calculating the R^2 value for the validation data set and comparing it to

the value computed for the estimation data set. Each of the two surveys conducted lasted one week. Data of the first five days of the week were used in estimation while data of the last two days were used in validation.

Table 4.2 Mathematical representation of the models.

Model no.	Survey no.	Output	Inputs	Model form and parameter estimates
M-1	1	Effluent TSS (Y_t)	Influent flow ($X_{1,t}$) and influent TSS ($X_{2,t}$)	$Y_t = \frac{\omega_{1,0}}{1} X_{1,t-1} + \frac{\omega_{2,0}}{1} X_{2,t-1} + \frac{1}{(1-\phi_1 B)} a_t$ $\omega_{1,0} = 0.222 (0.022)$ $\omega_{2,0} = 0.034 (0.015)$ $\phi_1 = 0.716 (0.07)$
M-2	2	Effluent TSS (Y_t)	Influent flow ($X_{1,t}$) and influent TSS ($X_{2,t}$)	$Y_t = \frac{\omega_{1,0}}{1} X_{1,t-1} + \frac{\omega_{2,0}}{1} X_{2,t-1} + \frac{1}{(1-\phi_1 B)} a_t$ $\omega_{1,0} = 0.173 (0.025)$ $\omega_{2,0} = 0.053 (0.024)$ $\phi_1 = 0.648 (0.072)$
M-3	2	Effluent COD (Y_t)	Influent flow ($X_{1,t}$) and influent COD ($X_{2,t}$)	$Y_t = \frac{\omega_{1,0}}{1} X_{1,t-1} + \frac{\omega_{2,0}}{(1-\delta_{2,1} B)} X_{2,t-1}$ $+ \frac{1}{(1-\phi_1 B - \phi_2 B^2)} a_t$ $\omega_{1,0} = -0.135 (0.046)$ $\omega_{2,0} = 0.170 (0.024)$ $\delta_{2,1} = 0.757 (0.036)$ $\phi_1 = 0.648 (0.097) \quad \phi_2 = 0.182 (0.096)$

Number in parentheses indicates standard error.

4.5.1 Survey #1

The objective was to build a transfer-function noise model that links the effluent TSS, denoted by Y_t , with the influent flow rate, denoted by $X_{1,t}$, and the influent TSS, denoted by $X_{2,t}$. This model is denoted by M-1 in Table 4.2. As was mentioned previously, in order to identify a transfer function model component that links an input variable X_t to an output variable Y_t , the prewhitened cross correlation function between the two series has to be estimated. In order to do so, a stochastic model that can adequately transform the input series X_t into white noise was first identified, estimated, and validated. For the flow data, an ARIMA (2, 0, 0)x(0, 0, 1)₂₄ model was found to represent the data the best, and hence, was used to transform both the flow and effluent TSS series before estimating the cross correlation function between the two series, which is shown in the top left graph of Figure 4.4. The influent TSS data series was represented the best by an ARIMA (1, 0, 1)x(1, 0, 0), and hence, this linear filter was utilized to prewhiten both the influent and effluent TSS series before estimating the cross correlation function between the two series, which is shown in the top right graph of Figure 4.4.

Some transfer of input to output has been detected, as indicated by the significant spikes at lag one and two. It was clear from Figure 4.4 that the delay parameter, b , is one hour. Considering the 2.5 hour theoretical detention time for the tank, calculated based on the average flow rate recorded during the survey conducted (201 ML/d), having a delay parameter equal to one hour clearly indicates the presence of short circuiting in the tank. Theoretical residence time curves never exist in practice, especially for full-scale sedimentation basins, because ideal settling plug-flow conditions are never attained in practice due to the existence of hydraulic turbulence, short circuiting, and density currents (Tebbutt, 1979). Therefore, the actual detention time is likely to be less than the theoretical detention time calculated from the tank volume and the rate of flow.

In the present study two tracer studies (one at high flow and the other at low flow) for sedimentation tank #5 were conducted using water softening salt (brine) as a tracer. For both studies, a slug input of tracer was dumped at the influent channel and the

conductivity of the effluent wastewater from the tank was measured at the same point that was used for sampling the effluent TSS and COD. At the time of dumping the tracer, the wastewater inflow to PST 2 was recorded to be 255 and 125 ML/d for the first and second tracer study, respectively. The outcome of the first study is shown in Figure 4.11. Time zero on the horizontal axis represents the time at which the slug was dumped. It is evident from Figure 4.11 that the peak concentration reached the effluent sampling point approximately 50 minutes (0.83 hr) from the time of dumping the tracer. The flow to PST 2 of 255 ML/d at the time of dumping the tracer corresponds to a theoretical detention time of 1.8 hr. Although not shown here, the outcome of the second tracer study also indicated the presence of short-circuiting. These findings support using a delay parameter "b" of one hour in the transfer function component of the model.

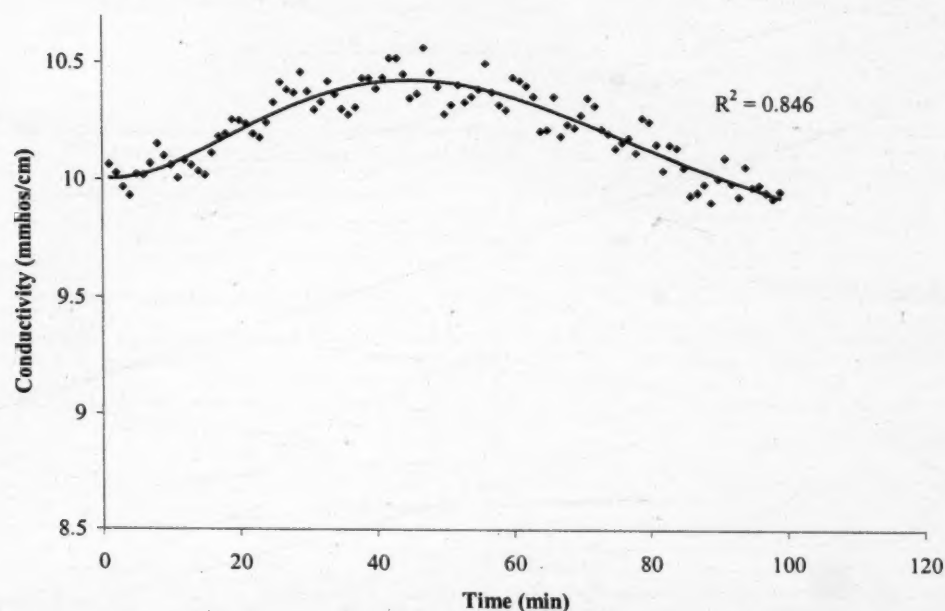


Figure 4.11 Outcome of the first tracer study conducted at high flow.

Although it was not possible to identify from Figure 4.4 whether the system behaves approximately according to some first-order transfer function, or whether a second-order model would be better, the cross correlation function indicated that a transfer function model with a numerator order of zero or one, and a denominator order of zero might be appropriate. Therefore, it was decided to fit several reasonable models

and select the best one that represents the data based on the diagnostic checks that were discussed earlier. It was found that the parsimonious model that best represented the data had a transfer function model component of order (0, 0, 1), that is both the numerator and denominator were of zero order and the delay parameter was equal to one time unit (one hour), and a noise model component of the form ARIMA (1, 0, 0). The equation for this model is shown in Table 4.2 along with values of the estimated parameters and their standard errors of estimate.

4.5.2 Survey #2

For the data that were collected during this survey, two transfer-function noise models were built in order to represent the dynamics of the primary sedimentation tank. The first model, denoted by M-2 in Table 4.2, links the effluent TSS, denoted by Y_t , with the influent flow rate, denoted by $X_{1,t}$, and the influent TSS, denoted by $X_{2,t}$. The second model, denoted by M-3 in Table 4.2, links the effluent COD, denoted by Y_t , with the influent flow rate, denoted by $X_{1,t}$, and the influent COD, denoted by $X_{2,t}$. An ARIMA (1,0,2) model was used in order to transform the input flow series into white noise before estimating the cross correlation function between it and the effluent series. The influent TSS series was prewhitened using an ARIMA (1, 0, 0) model. An ARIMA (1, 0, 1) model was used to transform the influent COD series into white noise.

After the prewhitening process, the estimated cross correlation functions were estimated and are shown in middle and bottom graphs of Figure 4.4 from which it is evident that the delay parameter is equal to one hour. For model M-2, it was not clear if a transfer function model with a numerator order of zero or one should be used, however, it was clear that a denominator order of zero should be used. It was found that the model that best represented the TSS data had a transfer function model component of order (0, 0, 1), and a noise model component of the form ARIMA (1, 0, 0). This model structure is identical to that of model M-1 developed for the TSS data of survey # 1. For model M-3, the cross correlation function between the transformed influent and effluent COD series clearly indicated a transfer function model with a denominator of first order. On the other

hand, the type of transfer function that links the flow data with the effluent COD data was not clear from the estimated cross correlation function between the two transformed series. It was found that the parsimonious model that best represented the COD data had a transfer function model component of order (0, 0, 1) relating the influent flow series to the effluent COD series while of order (1, 0, 1) relating the influent COD series to the effluent COD, and a noise model component of the form ARIMA (2, 0, 0). Table 4.2 shows the equations describing the models along with values of the estimated parameters and their standard errors of estimate.

4.5.3 Diagnostic Checking of the Models

As was mentioned previously, overfitting was used in order to test the validity of final models that were selected. In all instances, the 95% confidence limits associated with the extra (or overfit) parameters indicated that the additional parameters were not significantly different from zero. Additionally, there was little difference in the degree to which the overfitted models provided a better representation of the series being investigated.

For all of the three models M-1, M-2, and M-3, the standard errors for the parameter estimates, shown in parentheses in Table 4.2, indicated that the model parameters were significantly (with 95% confidence level) different from zero. Statistics calculated for the residuals as part of the diagnostic checks of the models are shown in Table 4.3. These statistics were calculated for the whole data set (including both the estimation and validation data sets), from which it is clear that for all of the three models, the mean error was not significantly (with 95% confidence) different from zero.

Residual diagnostics shown in Figure 4.5 to Figure 4.10 were performed on the whole data set. Figure 4.5 shows the autocorrelation and partial autocorrelation functions for the residuals from the models and indicates that both functions do not follow a specific pattern. Although few autocorrelations in Figure 4.5 appeared to be significantly different from zero, they were not clustered and were at high lags. In addition, they

hardly exceeded the confidence limits. A result that is significant in the statistical sense need not be important in the engineering sense, and therefore, these statistically significant spikes were felt to be unimportant from an engineering point of view.

Table 4.3 Model diagnostics.

Model ^a	ME ^b	S _a ^c	n ^d	$\frac{2 * S_a}{\sqrt{n}}$	MPE ^e	MAE ^f	MAPE ^g	R ^{2h}
M-1	0.18	8.93	168	1.38	-2.97	6.58	15.00	0.80
M-2	-1.02	12.39	168	1.91	-7.70	8.21	18.94	0.73
M-3	1.54	21.73	167	3.36	0.17	15.43	3.47	0.90

^aFor description of the model, see Table 4.2.

^bMean error (mg/L).

^cStandard deviation of the residuals (mg/L).

^dNumber of residuals.

^eMean percent error.

^fMean absolute error (mg/L).

^gMean absolute percent error.

^h $R^2 = 1 - \frac{\sum (a_i)^2}{\sum (Y_i - \mu)^2}$ where μ is the mean of the original series values Y_i .

Table 4.4 shows the Ljung-Box white noise test for residuals and it is evident that it supports the serial independence of the residuals as a group. The cumulative periodograms for the residuals are shown in Figure 4.6, from which it is apparent that the points clustered closely about the theoretical line and there was no evidence of periodic characteristics buried in the residual series. In addition, the Kolmogorov-Smirnov white noise test accepts the null hypothesis that the residuals series represents white noise. Histograms and normal probability plots of the residuals are shown in Figure 4.7 and Figure 4.8, respectively, which clearly support the assumption of normality. Figure 4.9 shows plots of the residuals against predicted values. These plots show a random scatter

around zero. Plots of the residuals versus the input series are shown in Figure 4.10. In all instances, the residuals appear to be independent of the input series.

Table 4.4 Ljung-Box white noise test for residuals.

Model	Ljung-Box Q statistic ^b				
	Lag 6	Lag 12	Lag 18	Lag 24	Lag 30
M-1	1.10 (3)	7.05 (9)	20.25 (15)	40.19 (21)	46.89 (27)
M-2	4.95 (3)	13.10 (9)	19.19 (15)	30.57 (21)	35.34 (27)
M-3	2.19 (1)	7.31 (7)	12.88 (13)	17.23 (19)	20.29 (25)

^aFor description of the model, see Table 4.2.

^bNumber in parentheses indicates degrees of freedom.

As indicated by the value of R^2 shown in Table 4.3, model M-1, which is a combined transfer-function noise model, was able to account for approximately 80% of the variations within the effluent TSS data. Using only a noise component, an ARIMA (2, 0, 0)x(1, 0, 1)₂₄ was found to best fit the effluent TSS data of survey #1 and was able to account for 75% of the variations within the data. Using only a transfer function model component (without a stochastic component), that has the same structure of the transfer function component of model M-1, 60% of the variations within the data was accounted for. Approximately 73% of the variations within the effluent TSS data of survey #2 were accounted for by the model M-2. Using only a noise component, an ARIMA (1, 0, 0) was found to best fit the effluent TSS data of survey #2 and was able to account for 61% of the variations within the data. Using only a transfer function model component, 60% of the variations within the data was accounted for. Model M-3 accounted for approximately 90% of the variations within the effluent COD data of survey #2. Using only a noise component, an ARIMA (1, 0, 0) was found to best represent the data and was able to account for 86% of the variations within the data. Using only a transfer function model component, 82% of the variations within the COD data was accounted for.

4.5.4 Validating the Models

The one-step-ahead predictions of the models as well as the values for the R^2 computed for the estimation and validation data sets are shown in Figure 4.12 to Figure 4.14.

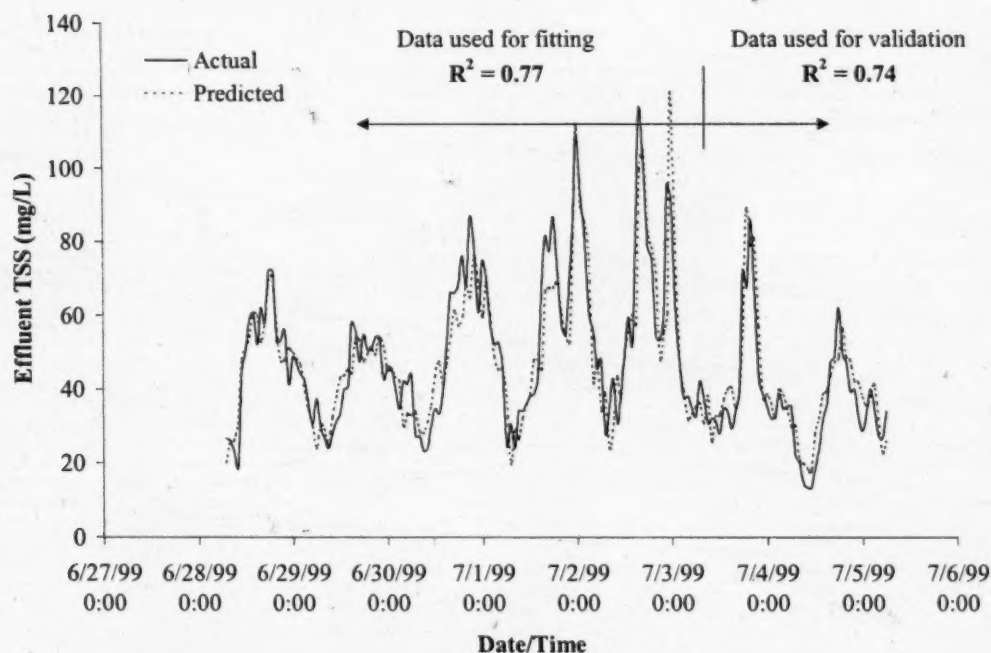


Figure 4.12 One-step-ahead forecasts by model M-1. For description of the models, see Table 4.2.

Even though models M-1 and M-2 (the TSS models) have the same structure, the accuracy of their forecasts for the validation data sets were different. Model M-1 gave an R^2 value of 0.74 for the validation data set, which was very close to the value of 0.77 obtained for the estimation data set. However, for model M-2, the R^2 value for the validation data set was almost half the value for the estimation data set. As it is clear from Figure 4.1, both the data sets used in estimating and validating model M-1 included rainfall events that had similar characteristics in terms of the flow measured during the event. Therefore, because of the fact that both of the two data sets included similar

features, model M-1 was able to generalize well when it was validated with the data set that was not seen by the model during the course of estimation.

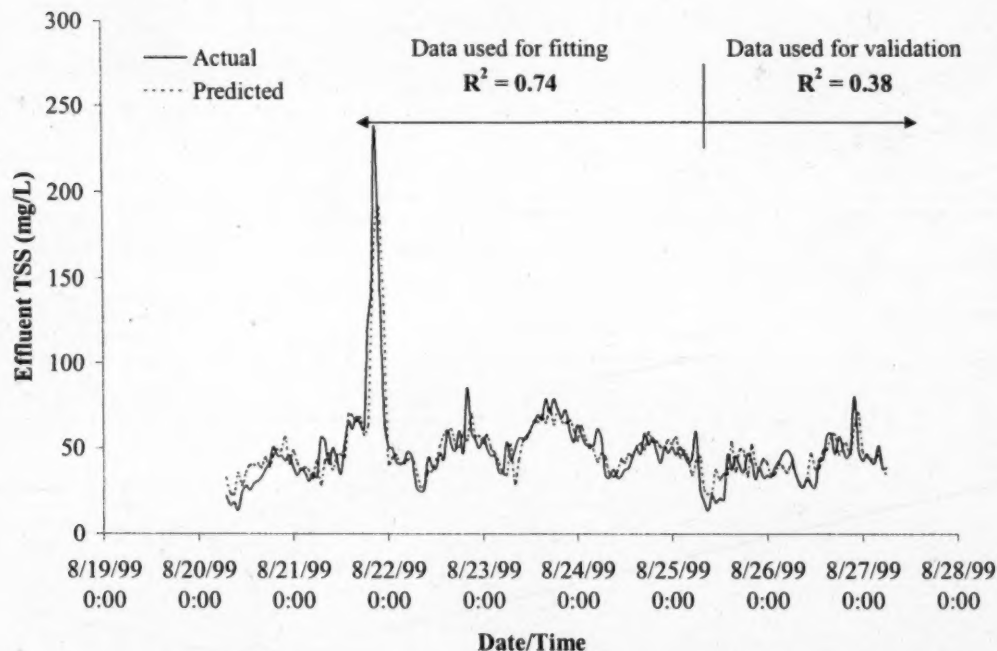


Figure 4.13 One-step-ahead forecasts by model M-2. For description of the models, see Table 4.2.

On the other hand, during survey # 2, only one rainfall event took place (labeled "rain event # 6" in Figure 4.2). This Event, which commenced on August 21 at 17:00 hours, had the highest intensity among the rainfall events that were encountered during the sampling periods conducted in the study. During this event the flow increased by more than 200% of the normal dry weather values. Because of the fact that this event was included in the estimation data set for model M-2, the values for the model parameters estimated were biased toward fitting the data points of this extreme event. Therefore, when the estimated model was tested against the verification data set, the accuracy of the forecasts was dramatically reduced as indicated by the value of the R^2 . The accuracy of the forecasts could have been improved if the sampling period would last more than one week in order to collect more data, especially, during rainfall events. However, this was

not possible due to labor limitations. The COD model M-3 gave an R^2 value of 0.84 for the validation data set, which was very close to the value of 0.88 obtained for the estimation data set.

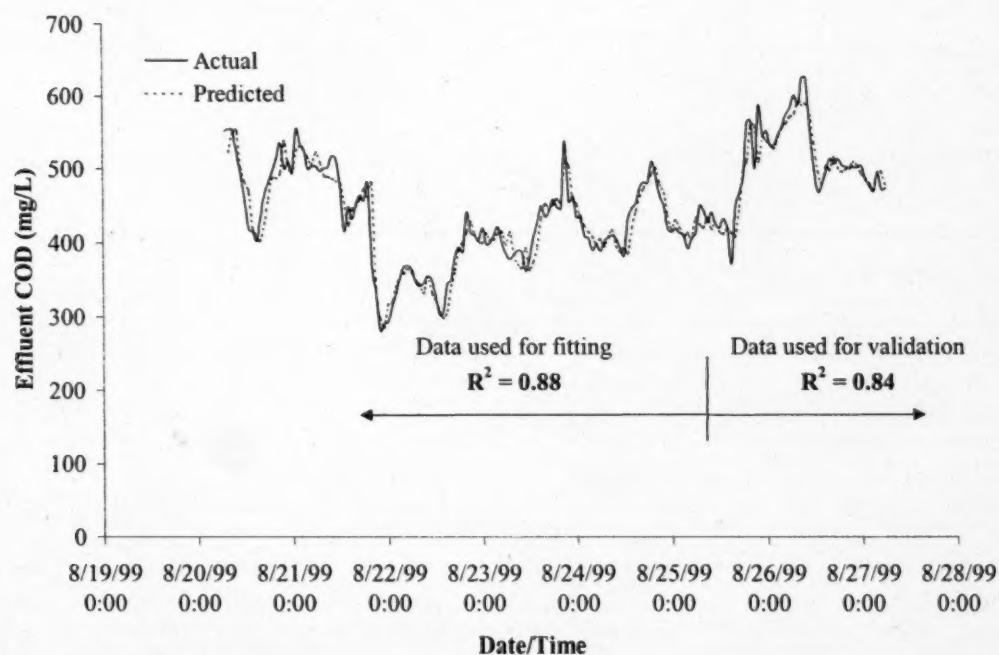


Figure 4.14 One-step-ahead forecasts by model M-3. For description of the models, see Table 4.2.

4.6 POSSIBLE APPLICATIONS OF THE MODELS

The primary sedimentation process at the GBWWTP is followed by biological treatment in a suspended growth activated sludge system, final settling, and microorganism reduction. However, during many of the rainfall events the capacity of the secondary treatment is exceeded and secondary bypass is utilized. During such events, real-time control of the primary sedimentation process would be of a great value in order to minimize the pollutional impact on the receiving water. The stochastic models described in the present study may be integrated into such a control scheme. The effluent TSS and COD would be the output variables that need to be targeted at a certain "target

value", which may be estimated using water quality modelling of the receiving stream. A combined feedforward-feedback control scheme may be implemented for the primary sedimentation process. The effect of measured uncontrolled sources of disturbance (such as influent TSS and COD), represented by the transfer function components that relate influent TSS and COD to effluent TSS and COD, may be accounted for by feedforward control. Only the effect of unmeasured sources of disturbance (such as hydraulic turbulence, density currents, short circuiting, measurement errors, etc.) on the output, represented by the "noise" process N_t in Equation 4.7, may be accounted for by feedback control. In such control scheme, the influent flow may be used as the measurable controlled variable that is utilized in order to bring the output variable back to its target value. Controlling the flow entering a primary sedimentation process may be implemented by means of flow equalization and/or flow redistribution (bringing one or more out-of-service tanks back to service during rainfall events).

In order to implement such a control scheme in reality, online measurements of the flow and quality parameters (TSS and COD) are needed. The flow volume is measured online at the GBWWTP and can be incorporated into such a use. On the other hand, it is impossible to measure the TSS online. However, the quality of wastewater with respect to colloidal and suspended matter may be measured by turbidity. Linear relationships between total suspended solids and turbidity of wastewater treated by an activated sludge system have been established. Although this type of relationships for untreated wastewater is much more difficult to establish, some researchers (Vanderborght and Wollast, 1990; Bertrand-Krajewski, 1992) developed such relationships for the purpose of monitoring the total suspended solids on real-time basis. Online COD analyzers are available, however, they are expensive and therefore more frequently used in the control of biological processes. They utilize ozone for the oxidation of the organic matter. Londong and Wachtl (1996) used a COD analyzer to monitor the influent untreated wastewater to a treatment plant in Germany. Instead of monitoring the COD itself, other parameters that can be correlated to the influent COD, and in the same time can be measured online, may be utilized. Among these parameters are the conductivity and turbidity. They have the ability to give information about the dissolved and suspended

load in the wastewater flow. Hack and Kohne (1996) found very strong correlation between each of the conductivity and turbidity and the influent COD to a treatment plant.

The development of such a control scheme should be iterative. Using operating data such as the ones collected in the current study, preliminary transfer function and noise models are postulated and used to design a pilot control scheme. The operation of this pilot scheme can then be used to supply further data. As additional data for a time series are made available, the same model for the original series can be re-estimated and then used to generate new (revised) forecasts based on this additional data by moving the forecast origin forward in time with the length of the additional data. When the model is updated, either the current model is re-estimated by using the new data or a change is made to the model structure, i.e., a new model must be identified and estimated. However, since it is unlikely that the basic relationships that existed in the original series will change drastically because of the new data, a new model will seldom need to be identified. In the current study the TSS models (M-1 and M-2) had exactly the same structure but the estimated values of the parameters were different which reflected the different operating conditions (such as temperature) during the two surveys conducted (one in May-June of 1999 and the other in August of 1999). The models described here are highly adaptive. In other words, they can be updated (re-estimated) on regular basis with minimal efforts.

4.7 CONCLUSIONS

Understanding and modelling the dynamics of a process is one of the essential steps towards designing a control scheme for that particular process. The objective of this study was to study the dynamics of a full-scale primary sedimentation tank using combined transfer-function noise models. The methods reported here represent a way by which plant data speak for themselves about the dynamics of the process. The procedure for identifying the models was described. A comprehensive system of diagnostic checks was utilized in order to validate the models. It was possible to build stochastic transfer function models which describe the data well. These models accounted for approximately

76% and 90% of the total variation in the primary effluent TSS and COD response series, respectively. This is judged satisfactory considering normal measurement errors.

The relative importance of the two components, namely the transfer function model component and the stochastic (noise) component, comprising the models has been assessed. These results showed that the stochastic part of the model is extremely important, especially for the modelling of the TSS data. It was also evident that the transfer function component between the influent and effluent COD data was more pronounced than that between the influent and effluent TSS data. In other words, the primary sedimentation process had the ability to dampen out the variations of the influent TSS more effectively than its ability to dampen out the variations of the influent COD and this is due to the nature of the process itself. Dissolved and colloidal solids are not removed by the primary sedimentation process, and because a big part of the primary influent COD is in these forms, the variability in the influent TSS is more effectively dampened out by the primary sedimentation process than that of the influent COD.

With respect to the hydraulics of the tank, it was found that the tank suffers from short-circuiting. Despite this fact, the increase in TSS load entering the tank during rainfall events, that were encountered during the surveys conducted, was dampened out by the primary sedimentation process. In the plant studied in this project, there is no process control applied to primary sedimentation. The findings of the present study suggest that with the current operational strategy implemented at the plant, during dry weather flow conditions, no real-time process control is needed for the primary sedimentation process. However, during rainfall events, during which the secondary capacity of the plant is exceeded, on-line process control of the primary sedimentation section would be valuable in order to minimize the pollutional load on the receiving stream. The present paper demonstrated the ability of the Box-Jenkins transfer function methodology to represent the stochastic dynamic nature of the primary sedimentation process and to provide short-term predictions of the quality data of the primary effluent wastewater.

4.8 REFERENCES

1. APHA-AWWA-WEF, *Standard Methods for the Examination of Water and Wastewater*, 19th Ed. American Public Health Association, American Water Works Association, and Water Environment Federation (Washington, DC, USA: American Public Health Association, 1995).
2. Bartlett, M.S., *Stochastic Processes* (Cambridge, UK: Cambridge University Press, 1955).
3. Berthouex, P.M., W.G. Hunter, and L. Pallesen, "Dynamic Behavior of an Activated Sludge Plant", *Wat. Res.* 12: 957-972 (1978).
4. Bertrand-Krajewski, J.L., "A Model for Solid Production and Transport for Small Urban Catchments: Preliminary Results", *Wat. Sci. Tech.* 25(8): 29-35 (1992).
5. Box, G.E. and G.M. Jenkins, *Time Series Analysis: Forecasting and Control* (Oakland, CA, USA: Holden-Day, 1976).
6. Capodaglio, A.G., "Evaluation of Modelling Techniques for Wastewater Treatment Plant Automation", *Wat. Sci. Tech* 30(2): 149-156 (1994).
7. Capodaglio, A.G., S. Zheng, V. Novotny, and X. Feng, "Stochastic System Identification of Sewer-Flow Models", *J. Env. Eng., ASCE* 116(EE2): 284-298 (1990).
8. Gamal El-Din, A. and Smith, D.W. "Modeling a Full-Scale Primary Sedimentation Tank Using Artificial Neural Networks", *Env. Tech.* 23: 479-496.
9. Hack, M. and M. Kohne, "Estimation of Wastewater Process Parameters Using Neural Networks", *Wat. Sci. Tech.* 33(1): 101-115 (1996).
10. Ljung, G.M. and G.E.P. Box, "On a Measure of Lack of Fit in Time Series Models", *Biometrika* 65:297-303 (1978).
11. Londong, J. and P. Wachtl, "Six Years of Experience with the Operation of On-Line Analyzers", *Wat. Sci. Tech.* 33(1): 159-164 (1996)
12. Novotny, V., A. Capodaglio, and H. Jones, "Real Time Control of Wastewater Treatment Operations", *Wat. Sci. Tech.* 25(4-5): 89-101 (1992).
13. Tebbutt, T.H.Y., "Primary Sedimentation of Wastewater", *J. WPCF* 51(12): 2858-2867 (1979).

14. Vanderborght, J.P. and P. Wollast, "Continuous Monitoring of Wastewater Composition in Sewers and Stormwater Overflows", Wat. Sci. Tech. 22(10/11): 271-275 (1990).

CHAPTER 5. DYNAMIC MODELLING OF A FULL-SCALE ACTIVATED SLUDGE PROCESS USING A SIMPLIFIED VERSION OF THE IWA-ASM1*

5.1 INTRODUCTION

In general, the activated sludge process is a continuous or semi-continuous (fill-and-draw) aerobic method for biological wastewater treatment (Ganczarczyk, 1983). The first activated sludge treatment plant was put into operation in England in 1916. The activated sludge process is still one of the most widely utilized secondary treatment processes for treating municipal and industrial wastewaters. The main function of the process is the removal of biodegradable organics and the production of an effluent, which is low in both degradable organics and suspended solids. Both nitrogen and phosphorus can impact receiving water quality. As a result, research has determined several methods for nutrient removal from wastewater prior to disposal. One method of removal is the biological nutrient removal (BNR) method. BNR processes are modifications of the activated sludge process that incorporate anoxic and/or anaerobic zones to encourage nitrogen and/or phosphorus removal.

The modelling of biological wastewater treatment systems has passed through a sequence of events: first, the removal of organic matter only; second, nitrification; third, nitrogen removal by biological denitrification; and forth, phosphorus removal by biological treatment of wastewater. Monod (1949) described the different growth phases of a bacterial culture and introduced the kinetics that describe the relation between the exponential growth rate of bacteria and the concentration of a limiting nutrient; an empirical relationship that was the origin of mathematical models of continuous growth microbial systems. Examples of these models that describe the activated sludge process are McKinney (1962), Lawrence and McCarthy (1970), and Eckenfelder (1985).

Realizing the benefits to be drawn from mathematical modelling, the International Water Association (IWA; formerly the International Association on Water Quality IAWQ) formed a task group in 1983 to promote the development of practical models to

A version of this chapter has been submitted for publication. Gamal El-Din, A. and D.W. Smith, *Journal of Environmental Engineering and Science* (May 2002).

the design and operation of biological wastewater treatment systems (Henze et al., 1987). Reviewing existing models was the first goal of the task group and the second goal was to reach a consensus concerning the simplest model able to predict the performance of an activated sludge system carrying out carbon oxidation, nitrification and denitrification (Henze et al., 1987). Activated Sludge Model No. 1 (ASM1) was the model developed by the IWA task group in 1987. ASM1 has proved to be an excellent tool for modelling nitrification-denitrification processes (Barker and Dold, 1997, Daigger and Nolasco, 1995, Henze et al., 1995 and Zhao et al., 1997).

In order to model biological nutrient removal in complex activated sludge processes, the model structure requires a high dimension and the model possesses a large number of stoichiometric and kinetic parameters. As a result, the models have grown more complex over the years, from ASM1, including nitrogen removal processes, to ASM2, released in 1995 and including biological phosphorus removal processes, and to ASM2d released in 2000 and including denitrifying Phosphorus Accumulating Organisms (PAOs). Scientific research and model application in engineering practice have different goals. Whereas the detailed structure of the models is used in order to describe new mechanisms which have been identified in advanced research projects, manageable models with a moderate number of parameters but a high potential to predict system behavior should be the basis for model application in engineering practice (Henze et al., 2000). A simplified version of ASM1 is utilized in the current study in order to model a full-scale activated sludge tank at the Gold Bar Wastewater Treatment Plant (GBWWTP), the largest plant in the Edmonton area.

During the months of July and August of every year the GBWWTP, usually suffers from settling problems occurring in the secondary clarifier. In order to gain insight into the process, an extensive sampling campaign was conducted on one of the activated sludge tanks at the plant during the summer of 2000 and as part of this campaign two sampling surveys were conducted, one in May/June and the other in July/August. Dynamic modelling of the activated sludge process was conducted using the experimental data that were collected. Description of the study, model development,

results of the modelling effort as well as conclusions and further research objectives are presented in the following sections.

5.2 STUDY

5.2.1 The Gold Bar Wastewater Treatment Plant (GBWWTP)

The GBWWTP was constructed in 1956 on the southwest shore of the North Saskatchewan River. The initial dry-weather flow capacity of the plant was 91 ML/d, serving the needs of a population of 250,000. The plant was expanded in 1971-72 and in 1979-81. The present capacity of the plant is 950 ML/d for primary treatment and 420 ML/d for secondary treatment based on both hydraulic and process capacities. Approximately 95% of the sewage flows from the City of Edmonton are treated at the GBWWTP. The Gold Bar Wastewater Treatment Plant is typical of many conventional activated sludge plant designed for carbonaceous BOD and suspended solids removal. It provides both primary and secondary treatment for the incoming raw sewage. Primary treatment consists of grit removal, mechanical screening, and primary sedimentation. The secondary treatment provides biological treatment in a suspended growth activated sludge system, final settling and microorganism reduction.

5.2.2 The Activated Sludge Process at the GBWWTP

At the GBWWTP there are ten rectangular aeration basins which are numbered 1 through 10. Each aeration basin has a corresponding rectangular secondary clarifier. Listed in Table 5.1 are the physical dimensions of the activated sludge tanks. In the current study, activated sludge tank #8 was sampled. The tank is part of a unique configuration at the GBWWTP that is composed of activated sludge tanks #6, 7 and 8. This configuration is shown in Figure 5.1. The three tanks are identical in dimensions and configuration. The mixed liquor from the three aeration basins is combined in a distribution chamber that distribute the mixed liquor among the three secondary clarifiers #6, 7 and 8. The return activated sludge (RAS) from each of the three clarifiers is combined in a distribution chamber that divides the RAS among the three aeration tanks.

The reason that aeration tank #7 is in dotted line format in Figure 5.1 is that this bioreactor is normally out of service while the three secondary clarifiers are normally in service. This was the case in both of the sampling surveys that were conducted. The operating strategy at the GBWWTP is to control the value of the mixed liquor suspended solids (MLSS) in the aeration tank by changing the flow rate of the waste activated sludge (WAS) and keeping the flow rate of the RAS from each of the secondary clarifiers #6, 7 and 8 at the same value of approximately 20 ML/d. By controlling the value of the MLSS in the aeration tank, the operator can control the solids residence time (SRT) of the activated sludge process.

Table 5.1 Physical dimensions of the activated sludge tanks.

Physical dimension	Aeration basin	Secondary clarifier
Number of passes	4	N/A
Length (m)	97.5	76
Width (m)	6.2	26
Depth (m)	4.5	3.9
Tank Volume (m ³)	10570	7530

As shown in Figure 5.2, the aeration basin is divided into four longitudinal compartments or passes. Passes are connected at alternating ends. From the influent channel to the aeration basin, primary effluent flows northward in a Y wall channel between the second and third pass of each basin and at the north end of the Y wall channel is another channel running west that is used for discharging primary effluent into the beginning of the first pass. In addition to the first pass gate, set along each side of the Y wall channel are a number of gates for feeding primary effluent into the second and third pass of each basin. These gates are used in order to allow the operator to switch from plug flow mode to step-feed mode of operation. The normal strategy at the GBWWTP is to switch aeration tanks from plug flow mode (the normal mode of operation) to step-feed mode during the months of July and August of each year in order to accommodate the higher wastewater flow rates usually encountered during these two

months. The white thick arrows in Figure 5.2 represent the points of feed of the primary effluent to the aeration basin.

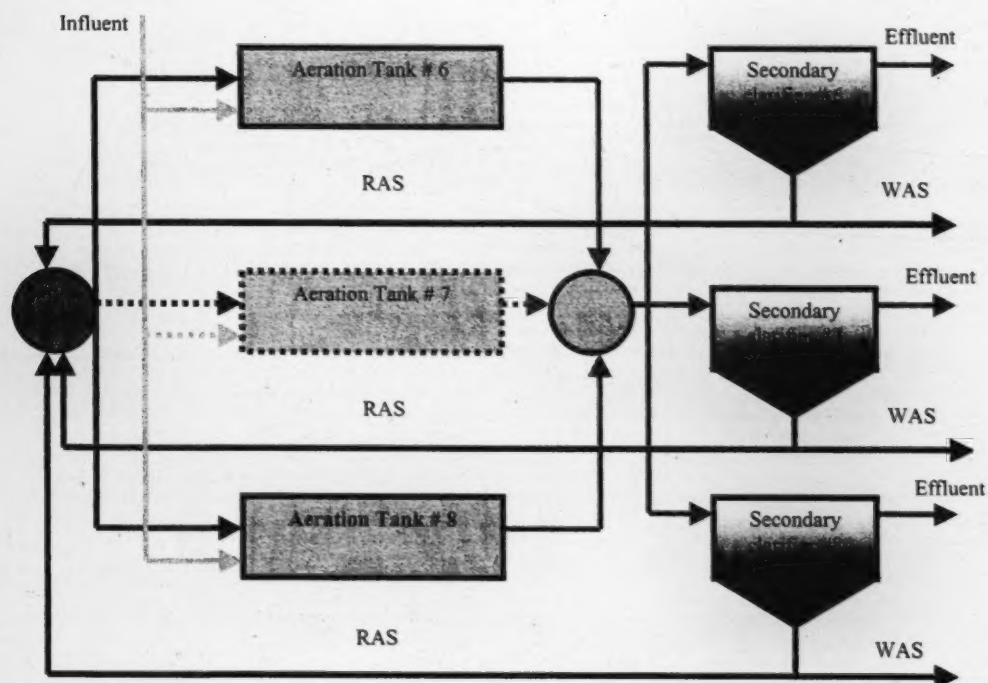


Figure 5.1 Activated sludge tanks #6, 7 and 8.

Coarse air bubble diffusers are used to supply air for mixing and oxygen transfer to the mixed liquor. Control of the aeration system is done through a combination of automatic and manual controls. At the north end of the second and fourth pass in the aeration basin are two dissolved oxygen (DO) probes. Through a modulating air valve, DO probes automatically control the amount of air supplied to maintain a specified DO at the probe location. The operating strategy at the GBWWTP is to control the dissolved oxygen concentration in the aeration basin at 2 mg/L at all times.

After mixing and aeration, mixed liquor is discharged into a short distribution channel running width wise across the front of the corresponding secondary clarifier. Air is supplied to the mixed liquor channels to keep the solids in suspension and to provide final process aeration prior to clarification. At Gold Bar, rectangular secondary clarifiers

are divided lengthwise into a number of cells. Each cell is equipped with a chain and flight collector and scum trough. Secondary Clarifiers #6, 7 and 8 each have four cells. Each cell has two manual inlet slide gates located at the south end for distributing mixed liquor into the tank. Just beyond the inlet slide gates is a wooden baffle wall. The baffle helps maintain quiescent conditions and reduces short circuiting in the clarifiers. Final effluent from each secondary clarifier flows over effluent weirs into common channels which combine together to direct the wastewater flow to the Ultra violet (UV) building where microorganism reduction of the secondary effluent takes place prior to discharge into the North Saskatchewan River.

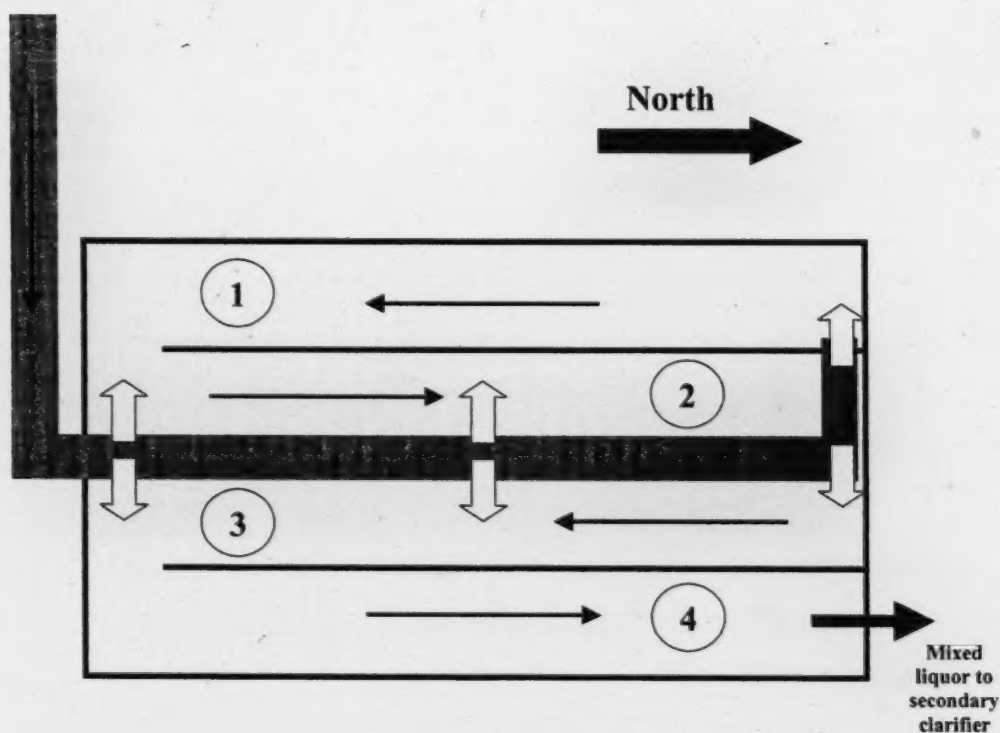


Figure 5.2 A schematic diagram of aeration tank #8 – Step-feed flow mode.

5.2.3 The Sampling Campaign

Samples were collected from both the primary effluent entering aeration tank #8 and the secondary effluent from secondary clarifier #8. Samples were also collected from

the mixed liquor in the 4th pass of the aeration tank and from the RAS line out of secondary clarifier #8. Two sampling surveys were conducted. The first survey started on May 23, 2001 at 9:00 hours and ended June 2, 2001 at 7:00 hours while the second survey started on July 28, 2001 at 9:00 hours and ended August 2, 2001 at 7:00 hours. During both of the sampling surveys, the wastewater flow was measured continuously at the effluent weir of each of the secondary clarifiers. Also both of the RAS flow rate and the waste activated sludge (WAS) flow rate from each of the secondary clarifiers #6, 7 and 8 were continuously monitored. Effluent wastewater, RAS, and WAS flow rate data were available for download using the Supervisory Control and Data Acquisition (SCADA) system at the plant.

In both surveys, bi-hourly samples were collected from both the primary effluent and the secondary effluent. Mixed liquor and RAS samples were collected every 6 hours. At 7:00 of every day during the sampling periods, the primary and secondary effluent samples of the previous 24 hour were taken to the laboratory at the University of Alberta. Each sample was divided into three portions. The first was used for total suspended solids (TSS) analysis. The second portion of the samples was preserved using concentrated sulfuric acid and stored in the cold room at 4°C to be analyzed later for chemical oxygen demand (total COD, COD_{TOT}), total Kjeldahl nitrogen (TKN) and total phosphorus (TP_{TOT}). The third portion of the samples was filtered using membrane filtration and the filtrate was preserved using concentrated sulfuric acid and stored in the cold room at 4°C to be analyzed later for chemical oxygen demand (soluble COD, COD_{SOL}), ammonia, nitrate and total phosphorus (Soluble TP, TP_{SOL}). At 7:00 hours of every day during the sampling periods, the mixed liquor and RAS samples of the previous 24 hour were taken to the laboratory at the GBWWTP and were analyzed for both TSS and volatile suspended solids (VSS). In both of the two surveys air temperature was recorded every two hours. Also, both the temperature and the pH of the primary effluent and mixed liquor were measured every two hours. All the analysis was conducted according to the Standard Methods for the Examination of Water and Wastewater (APHA, 1995).

5.2.3.1 The First Survey – Plug Flow Mode

This survey started on May 23, 2001 at 9:00 hours and ended June 2, 2001 at 7:00 hours. During this period all the activated sludge tanks at the plant were operated in the plug flow mode. During this survey the air temperature averaged 11.8°C while the temperature of the primary effluent averaged 16.6°C .

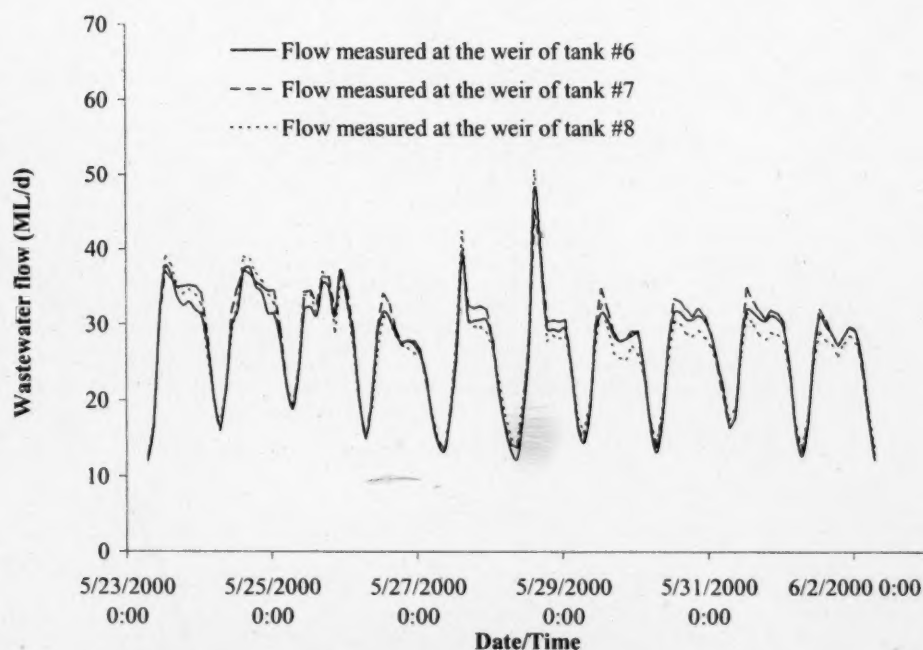


Figure 5.3 Wastewater flow measured at the effluent weir of each of the secondary clarifiers #6, 7 and 8 during the first survey.

Presented in Figure 5.3 are the data for the wastewater flow measured at the effluent weir of each of the secondary clarifiers #6, 7 and 8. It is evident from Figure 5.3 that the wastewater flow values measured for each of the three clarifiers are very close indicating that the flow is distributed approximately equally among the three tanks. The wastewater flow measured at the effluent weir of secondary clarifiers #6, 7 and 8 averaged 26.7, 27.4 and 26.6 ML/d, respectively. Two rainfall events took place during the 5th and 6th days of

the survey, as shown in Figure 5.3, during which the flow of wastewater measured at the effluent weirs of the clarifiers increased substantially above normal dry weather values.

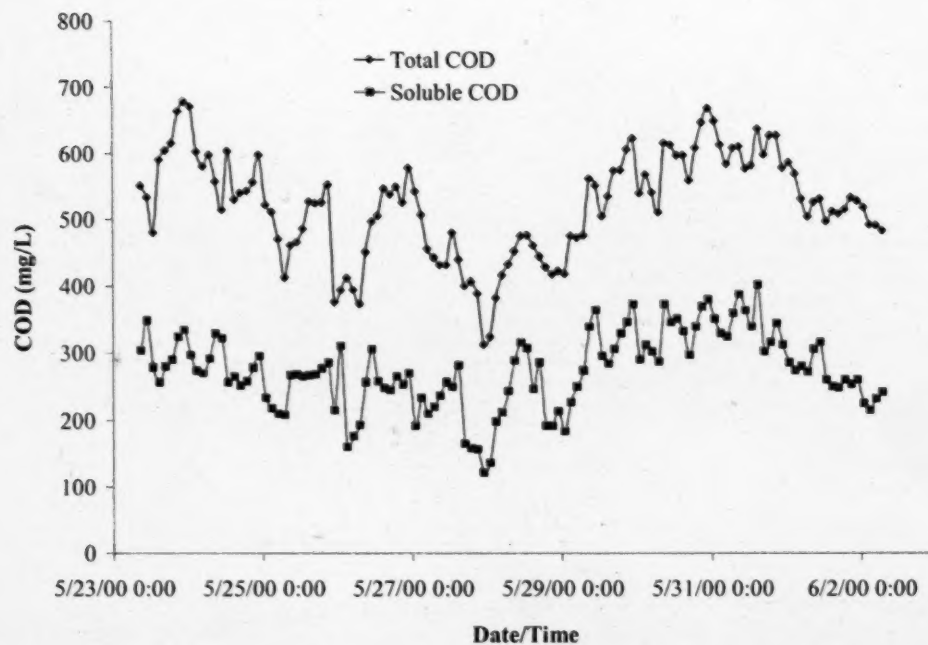


Figure 5.4 COD of the primary effluent samples collected during the first survey.

The TSS in the primary effluent averaged 95 mg/L. Figure 5.4 presents the COD data for the primary effluent samples collected from which it is evident that a weak diurnal pattern existed in the data. The total COD averaged 521 mg/L while the soluble COD averaged 274 mg/L. It is evident from Figure 5.3 and Figure 5.4 that during the two rainfall events that were encountered during the first sampling survey the COD concentrations decreased below average values.

The TKN and ammonia-nitrogen data for the primary effluent samples collected are shown in Figure 5.5. The TKN in the primary effluent averaged 40.5 mg/L-N while the ammonia-nitrogen averaged 29.5 mg/L-N. It is clear from Figure 5.5 that both the TKN and ammonia-nitrogen in the primary effluent followed the same general trend. A strong diurnal pattern existed in the data presented in Figure 5.5.

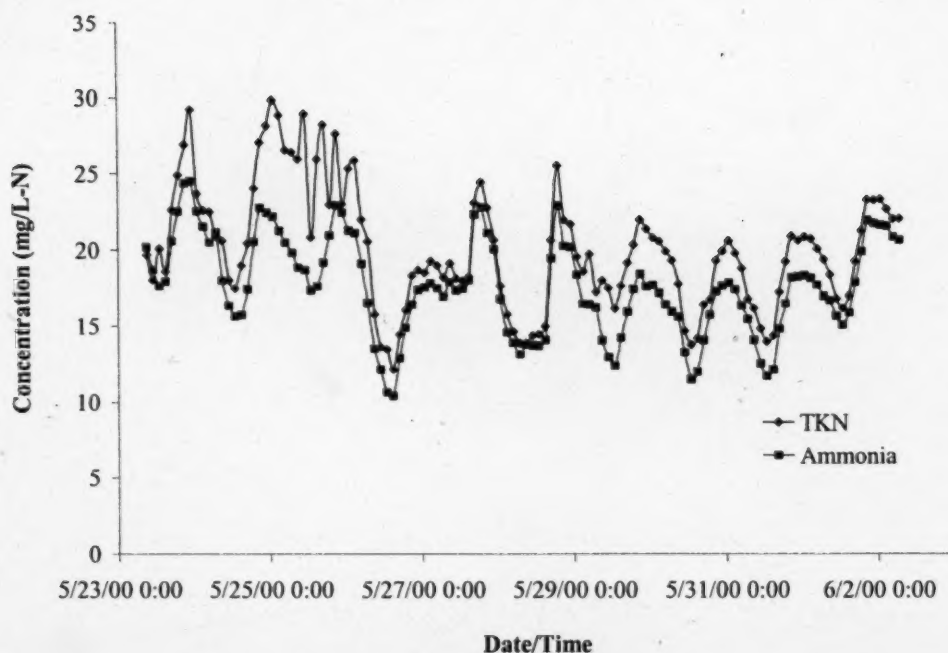


Figure 5.5 TKN and ammonia nitrogen concentrations in the primary effluent samples collected during the first survey.

5.2.3.2 The Second Survey – Step-Feed Mode

The second survey started on July 28, 2001 at 11:00 hours and ended August 4, 2001 at 7:00 hours. The step-feed mode was used for operating aeration tanks #6, 7 and 8 during this survey. During this survey the air temperature averaged 20.9°C while the temperature of the primary effluent averaged 20.5°C . These temperatures are higher than those recorded during the first sampling survey. As was the case in the first sampling survey, the wastewater flow data measured at the effluent weir of each of the secondary clarifiers #6, 7 and 8 were very close indicating equal flow distribution. The wastewater flow measured at the effluent weir of secondary clarifiers #6, 7 and 8 averaged 30.8, 32.4 and 30.8 ML/d, respectively. These average flow values are higher than that encountered during the first sampling survey.

The TSS in the primary effluent averaged 74.7 mg/L. It should be noted that TSS data for the primary effluent were available for only six days of the 7-days sampling survey. The total COD averaged 407 mg/L while the soluble COD averaged 220 mg/L. These values are lower than the average values for the first survey. Both the total and soluble COD in the primary effluent followed the same general trend. As was the case in the first survey, a weak diurnal pattern existed in the primary effluent COD data. The TKN in the primary effluent averaged 39.5 mg/L-N while the ammonia-nitrogen averaged 29 mg/L-N. These values are very close to the average values for the first survey. As was the case in the first survey conducted, both the TKN and ammonia-nitrogen in the primary effluent followed the same general trend and a strong diurnal pattern existed in the data.

5.2.3.3 Comparison between the Two Surveys

Table 5.2 Average values for the parameters measured for the primary and secondary effluents.

Parameter	Units	First survey	Second survey	First survey	Second survey
		Primary effluent		Secondary effluent	
Total COD	mg/L	521.5	406.7	70	54.5
Soluble COD	mg/L	273.9	219.8	35.3	46.4
TKN	mg/L	40.5	39.5	20	14.1
Ammonia-N	mg/L-N	29.5	29	17.6	12.5
Nitrate-N	mg/L-N	0.1	0	0.4	4.3
Total P	mg/L-P	5.2	5.6	0.8	2.5
Soluble P	mg/L-P	4.1	3.9	0.6	2.4
pH		7.3	7.2	NM ^a	NM ^a
Temperature	°C	16.6	20.5	NM ^a	NM ^a
TSS	mg/L	94.8	74.7	9.8	2.4

^aNot measured

Average values for the parameters measured for both the primary and secondary effluent from activated sludge tank #8 are listed in Table 5.2. Table 5.3 lists the average values for the parameters measured for the mixed liquor samples from the 4th pass and the RAS samples. While the ammonia-N concentration in the secondary effluent averaged 17.6 mg/L during the first sampling survey, it only averaged 12.5 mg/L during the second survey. This was because partial nitrification took place during the second survey. Although nitrification in the group of activated sludge tanks 6,7 and 8 was not one of the treatment objectives at the GBWWTP, partial nitrification took place in the second survey due to the elevated water temperature of the mixed liquor in the aeration tank. Examining Table 5.3 reveals that the water temperature of the mixed liquor averaged 21°C during the second survey while it only averaged 16.9°C during the first one. The nitrate-N concentrations shown in Table 5.2 also clearly indicate that nitrification took place in the aeration tank during the second survey while only very slight nitrification occurred during the first survey.

Table 5.3 Average values for the parameters measured for the mixed liquor and return activated sludge.

Parameter	Units	First survey	Second survey	First survey	Second survey
		Mixed liquor from the 4 th pass		Return activated sludge	
pH		7.1	7.1	NM ^a	NM ^a
Temperature	°C	16.9	21	NM ^a	NM ^a
TSS	mg/L	1808	1263	4196	3171
VSS	mg/L	1479	1050	3386	2566

^aNot measured

It will be shown later that the modelling results clearly indicate that denitrification was taking place in the secondary clarifiers during the second survey. This phenomenon was also observed during the settling tests that were conducted by the plant operators for the mixed liquor from the 4th pass during the period of the second survey. After

approximately 30 minutes from the beginning of a test, sludge started to rise which was attributed to the release of nitrogen gas. Besides the change in the operational mode from plug flow in the first survey to step-feed in the second, the partial nitrification-denitrification that took place during the second survey were the main differences between the conditions encountered during the first and the second surveys.

5.3 MODEL DEVELOPMENT

The model used in the current study is a simplified version of the Activated Sludge Model No. 1 (ASM1) developed by the International Water Association (IWA). ASM1 is able to predict the performance of an activated sludge system carrying out carbon oxidation, nitrification and denitrification (Henze et al., 1987). In the following sections of this paper, this simplified version of ASM1 will be called "ASM1_{SP}".

ASM1_{SP} has ten soluble and particulate components and six different processes. The model is presented in the matrix format in Table 5.4. This format for model presentation makes it easier to trace all the interactions of the different components in the model (Henze et al., 2000). Listed across the top of Table 5.4 by symbol are the different components in the model. The index i is assigned to each component. In this case, i ranges from 1 to 10 for the ten compounds in the model. Listed in Table 5.5 are the symbols, names and units of the different components in ASM1_{SP}. In conformity with IWA nomenclature, insoluble constituents are given the symbol X and the soluble components S . Subscripts are used to specify individual components: B for biomass, S for substrate and O for oxygen.

The ASM1_{SP} contains six biological processes which are conversions or transformations that affect the different components listed in Table 5.5. These processes are listed in the leftmost column of Table 5.4. The index j is assigned to each process; in this case, j ranges from 1 to 6 for the six different biological processes in the model. Recorded in the rightmost column of Table 5.4 in the appropriate row are the kinetic expressions or rate equations for each process.

Table 5.4 Process kinetics and stoichiometry for ASM_{SP} .

Component ^a i →	1	2	3	4	5	6	7	8	9	10
j ↓ Process	X _S	X _{B,H}	X _{B,A}	X _I	X _P	S _I	S _S	S _O	S _{NO}	S _{NH}
	Process Rate ρ _j , ML ⁻³ T ⁻¹									
1 Aerobic growth of heterotrophs		1					$-\frac{1}{Y_H}$	$\frac{1 - Y_H}{Y_H}$	$-i_{N/XB}$	$\hat{\mu}_H \left(\frac{S_S}{K_S + S_S} \right) \left(\frac{S_O}{K_{O,H} + S_O} \right) X_{B,H}$
2 Anoxic growth of heterotrophs		1					$-\frac{1}{Y_H}$	$-\frac{1 - Y_H}{2.86 Y_H}$	$-i_{N/XB}$	$\hat{\mu}_H \left(\frac{S_S}{K_S + S_S} \right) \left(\frac{K_{O,H}}{K_{O,H} + S_O} \right)$ $\left(\frac{S_{NO}}{K_{NO} + S_{NO}} \right) \eta_g X_{B,H}$
3 Aerobic growth of autotrophs			1					$\frac{4.57 - Y_A}{Y_A}$	$\frac{1}{Y_A}$	$-\frac{1}{i_{N/XB} - Y_A}$ $\hat{\mu}_A \left(\frac{S_{NH}}{K_{NH} + S_{NH}} \right) \left(\frac{S_O}{K_{O,A} + S_O} \right) X_{B,A}$
4 Death and lysis of heterotrophs					f _p					b _{L,H} X _{B,H}
5 Death and lysis of autotrophs					f _p					b _{L,A} X _{B,A}
6 Hydrolysis of particulate organics							1			$K_b \frac{X_S / X_{B,H}}{K_X + (X_S / X_{B,H})} X_{B,H}$
Observed conversion rates, ML ⁻³ T ⁻¹										

$$r_i = \sum_{j=1}^n \Psi_{ij} r_j$$

^aAll organic compounds (1 to 7) and oxygen (8) are expressed as COD; all nitrogenous components (9 to 10) are expressed as nitrogen.

Table 5.5 Different components in ASM1_{SP}.

Component Symbol	Component Name	Units
X_S	Slowly biodegradable substrate	$[M(COD)L^{-3}]$
$X_{B,H}$	Active heterotrophic biomass	$[M(COD)L^{-3}]$
$X_{B,A}$	Active autotrophic biomass	$[M(COD)L^{-3}]$
X_I	Particulate inert organic matter	$[M(COD)L^{-3}]$
X_P	Particulate product arising from biomass decay	$[M(COD)L^{-3}]$
S_I	Soluble inert organic matter	$[M(COD)L^{-3}]$
S_S	Readily biodegradable substrate	$[M(COD)L^{-3}]$
S_O	Oxygen (negative COD)	$[M(-COD)L^{-3}]$
S_{NO}	Nitrate and nitrite	$[M(N)L^{-3}]$
S_{NH}	$NH_4^+ + NH_3$ nitrogen	$[M(N)L^{-3}]$

The kinetic parameters in ASM1_{SP} are defined in Table 5.6. Process rates are denoted by p_j in Table 5.4 where j corresponds to the process as numbered in the leftmost column of the table. The elements within the matrix comprise the stoichiometric coefficients, v_{ij} , which set out the mass relationships between the components in the individual processes. The coefficients, v_{ij} , are greatly simplified by working in consistent units. In this case, all organic constituents have been expressed as equivalent amounts of chemical oxygen demand (COD); likewise, oxygen is expressed as negative oxygen demand. The sign convention used in the matrix is negative for consumption and positive for production. All stoichiometric coefficients are defined in Table 5.7.

The concentration of a single component within a system may be affected by a number of different processes. The matrix representation of the model allows easy recognition of the fate of each component, which aids in the preparation of mass balance equations (Henze et al., 2000). The basic equation for a mass balance within any defined system boundary is:

$$\text{Input} - \text{Output} + \text{Reaction} = \text{Accumulation} \quad [5.1]$$

The input and output terms are transport terms that depend on the physical characteristics of the system being modelled. The system reaction term, r_i , for component i is obtained by moving down the column representing the component and summing the products of the stoichiometric coefficients v_{ij} and the process rate expression ρ_j for the component i being considered in the mass balance:

$$r_i = \sum_j v_{ij} \rho_j \quad [5.2]$$

For example the rate of reaction for heterotrophic biomass, $X_{B,H}$, at a point in the system would be:

$$r_{X_{B,H}} = \hat{\mu}_H \left(\frac{S_S}{K_S + S_S} \right) \left(\frac{S_O}{K_{O,H} + S_O} \right) X_{B,H} + \hat{\mu}_H \left(\frac{\hat{S}_S}{K_S + \hat{S}_S} \right) \left(\frac{K_{O,H}}{K_{O,H} + S_O} \right) \left(\frac{S_{NO}}{K_{NO} + S_{NO}} \right) \eta_g X_{B,H} - b_{L,H} X_{B,H} \quad [5.3]$$

Another benefit of the matrix presentation is that continuity may be checked by moving across the matrix, provided consistent units have been used because then the sum of the stoichiometric coefficients must be zero. For example, considering the death and lysis process for the heterotrophic biomass, two continuity checks, one for COD and the other for nitrogen, can be done. The COD continuity check for the death and lysis of heterotrophic biomass would be:

$$(1 - f_p) * 1 + (-1) * 1 + f_p * 1 + (i_{N/XB} - f_p i_{N/XP}) * 0 = 0 \quad [5.4]$$

The nitrogen continuity check for the death and lysis of heterotrophic biomass would be:

$$(1 - f_p) * 0 + (-1) * i_{N/XB} + f_p * i_{N/XP} + (i_{N/XB} - f_p i_{N/XP}) * 1 = 0 \quad [5.5]$$

Table 5.6 Kinetic parameters in ASM1_{SP} (Grady et al., 1999).

Symbol	Kinetic Parameter	Units	Value @ 20°C
$\hat{\mu}_H$	Maximum specific growth rate for heterotrophic biomass	hr ⁻¹	0.25
K_S	Half-saturation coefficient for heterotrophic biomass	mg/L as COD	20
$K_{O,H}$	Oxygen half-saturation coefficient for heterotrophic biomass	mg/L as O ₂	0.1
K_{NO}	Nitrate half-saturation coefficient for denitrifying heterotrophic biomass	mg/L as N	0.2
$b_{L,H}$	Decay coefficient for heterotrophic biomass	hr ⁻¹	0.017
η_g	Correction factor for heterotrophic growth under anoxic conditions	dimensionless	0.8
K_h	Maximum specific hydrolysis rate	mg COD(mg biomass COD.hr) ⁻¹	0.092
K_X	Half-saturation coefficient for hydrolysis of slowly biodegradable substrate	mg COD(mg biomass COD) ⁻¹	0.15
$\hat{\mu}_A$	Maximum specific growth rate for autotrophic biomass	hr ⁻¹	0.032
K_{NH}	Ammonia half-saturation coefficient for autotrophic biomass	mg/L as N	1
$K_{O,A}$	Oxygen half-saturation coefficient for autotrophic biomass	mg/L as O ₂	0.75
$b_{L,A}$	Decay coefficient for autotrophic biomass	hr ⁻¹	0.004

Table 5.7 Stoichiometric coefficients in ASM1_{SP} (Grady et al., 1999).

Symbol	Stoichiometric coefficient	Units	Value
Y_H	Heterotrophic yield	mg biomass COD formed/mg COD removed	0.6
f_p	Fraction of biomass yielding particulate products	mg debris COD/mg biomass COD	0.08
$i_{N/XB}$	Mass N/Mass COD in biomass	mg N/mg COD in active biomass	0.086
$i_{N/XP}$	Mass N/Mass COD in products from biomass	mg N/mg COD in biomass debris	0.06
Y_A	Autotrophic yield	mg biomass COD formed/mg N oxidized	0.24

Similar to ASM1, ASM1_{SP} is a complex model that contains many soluble and particulate components and as the number of continuous flow stirred tank reactors (CFSTRs) used to represent the aeration tank increases the degree of complexity increases and numerical techniques had to be used in order to solve the mass balance equations for the different constituents in the model. Several organizations have developed computer codes for solving the simultaneous mass balance equations for the constituents in ASM1 and ASM2, allowing their application to a variety of bioreactor configurations (Grady et al., 1999). One such code is ASIM (Activated Sludge SIMulation Program) written by the Swiss Federal Institute of Environmental Science and Technology. ASIM is a flexible modelling tool that implements both ASM1 and ASM2. It was used in the present study. The flexibility in ASIM allows the selection of the different constituents and processes to be modelled. This feature allowed us the freedom to choose the model structure to be used in ASM1_{SP}.

5.3.1 Conceptual Model

Similar to ASM1, ASM1_{SP} uses COD as the measure of the concentration of organic material in wastewater. COD provides a link between electron equivalents in the

organic substrate, the biomass and the oxygen utilized (Henze et al., 2000). Therefore, continuity checks and mass balances can be easily made in terms of COD. Consequently, the concentrations of all organic materials, including biomass are in COD units.

Based on biodegradability, the organic matter in ASM1_{SP} is subdivided into two categories: non-biodegradable organic matter and biodegradable organic matter. Non-biodegradable organic matter is biologically inert and passes through an activated sludge system unchanged in form. Depending on their physical state, inert organic matter are further divided into two fractions: soluble and particulate. Inert soluble organic matter, S_I , leaves the system at the same concentration that it enters. Inert particulate organic matter, X_I , becomes enmeshed in the activated sludge and is removed from the system through sludge wastage. In ASM1_{SP} biodegradable organic matter is divided into two fractions: readily biodegradable and slowly biodegradable. For purposes of modelling, the readily biodegradable material, S_S , is treated as if it were soluble, whereas the slowly biodegradable material, X_S , is treated as if it were particulate (Henze et al., 2000).

The readily biodegradable material consists of relatively simple molecules that may be taken in directly by heterotrophic bacteria and used for growth of new biomass, whereas the slowly biodegradable material, consisting of relatively complex molecules, must be acted upon extracellularly and converted into readily biodegradable substrate before it can be used (Henze et al., 1987). It is assumed that conversion of slowly biodegradable substrate into the readily biodegradable form (hydrolysis) involves no energy utilization and thus there is no utilization of electron acceptor associated with it.

In ASM1_{SP} decay of biomass is assumed to result in the conversion of biomass into slowly biodegradable substrate and particulate products, X_P , which are inert to further biological attack (Dold et al., 1980). The latter are similar in concept to the endogenous mass of McKinney and Ooten (1969) and act to reduce the viability of the mixed liquor suspended solids in a bioreactor. In ASM1_{SP} the ammonia nitrogen serves as the nitrogen supply for synthesis of heterotrophic biomass and as the energy supply for growth of autotrophic nitrifying bacteria. The autotrophic conversion of ammonia nitrogen to nitrate

nitrogen (nitrification) is considered to be a single step process which requires oxygen. The nitrate formed may serve as terminal electron acceptor for heterotrophic bacteria under anoxic conditions yielding nitrogen gas. Both ammonia nitrogen, S_{NH} , and nitrate nitrogen, S_{NO} , are expressed as nitrogen.

Similar to ASM1, ASM1_{SP} employs the concept of switching functions to turn process rate equations on and off as environmental conditions are changed which is necessary for processes that depend upon the type of electron acceptor present such as nitrification. Autotrophic bacteria are able to grow only under aerobic conditions and their rate of growth will fall to zero as the dissolved oxygen concentration approaches zero, regardless of the concentration of their energy yielding substrate (Henze et al., 1987). This is modelled in ASM1_{SP} by including a dissolved oxygen switch in the rate equation for nitrification. The oxygen switching function has the form:

$$\frac{S_o}{K_o + S_o} \quad [5.6]$$

where S_o is the concentration of dissolved oxygen. The selection of a small value for K_o means that the value of the switching function is near unity for moderate dissolved oxygen (DO) concentrations but decreases to zero as the DO concentration approaches zero (Henze et al., 1987). The fact that the function is mathematically continuous helps to eliminate problems of numerical instability which can occur during simulations with models which include rate equations that are switched on and off discontinuously (Henze et al., 1987). Processes which occur only when dissolved oxygen is absent, such as the anoxic growth of heterotrophic bacteria (denitrification), are modelled using a switching function of the form:

$$\frac{K_o}{K_o + S_o} \quad [5.7]$$

The selection of a small value for K_O means that the value of the switching function approaches zero for moderate dissolved oxygen (DO) concentrations but approaches unity as the DO concentration approaches zero.

5.3.2 Components in ASM1_{SP}

The components of the model are shown in the top of Table 5.4 and in Table 5.5. Soluble inert and particulate inert organic matter, S_I and X_I , ($i = 6$ and $i = 4$) are not involved in any conversion process, and therefore, their columns in Table 5.4 contain no stoichiometric coefficients. Nevertheless, they are included because they are important to the performance of the process (Henze et al., 1987). Soluble inert organic matter is part of the secondary effluent COD while particulate inert organic matter becomes part of the mixed liquor volatile suspended solids (MLVSS).

Moving down the $i = 7$ column, it can be seen that the readily biodegradable substrate, S_S , is removed by growth of heterotrophic bacteria under either aerobic or anoxic conditions and is formed by hydrolysis of particulate organic matter. The $i = 1$ column reveals that slowly biodegradable organic matter, X_S , is removed by the hydrolysis process and is formed by the death and lysis of both heterotrophic and autotrophic bacteria.

The biomass in the system are represented by columns $i = 2$ and 3, with $X_{B,H}$ denoting the heterotrophic biomass and $X_{B,A}$ denoting the autotrophic biomass. The heterotrophic biomass is removed by death and lysis and is formed by growth under either aerobic or anoxic conditions. The autotrophic biomass is destroyed through decay and is formed by growth under only aerobic conditions.

The $i = 5$ column represents the particulate products arising from biomass decay, X_P (McKinney and Ooten, 1969). According to ASM1_{SP}, it is formed by decay of both heterotrophic and autotrophic biomass, but is not destroyed. In actuality, this fraction of biomass is probably not completely inert to biological attack, however, its rate of

consumption is so low that it appears to be inert within the solids residence times (SRTs) normally encountered in activated sludge systems (Henze et al., 1987).

The sum of the five particulate terms: X_S , $X_{B,H}$, $X_{B,A}$, X_P and X_I is the volatile solids concentration in the activated sludge system, but in COD units. Therefore, an appropriate conversion factor can be applied to convert from COD units to volatile suspended solids units (Henze et al., 1987).

The dissolved oxygen concentration, S_O , is represented in column $i = 8$. It can be seen that the processes included in ASM1_{SP} only act to remove oxygen from solution and none is given for its addition or transfer as only biological processes are included in the matrix. Oxygen utilization is associated only with aerobic growth of the heterotrophic and autotrophic biomass and none is consumed by microbial decay. This approach differs from the more traditional approach (Grady et al., 1999). Decay is assumed to result in the release of slowly biodegradable substrate which is converted, through hydrolysis, to readily biodegradable substrate that is used for cell growth. Thus the oxygen utilization normally associated directly with decay is modelled in ASM1_{SP} as if it occurs indirectly from growth of new biomass on released substrate (Henze et al., 1987). This approach of modelling biomass decay is called the "lysis regrowth approach" (Grady et al., 1999). The fact that the heterotrophic yield is less than unity, so that the amount of new biomass grown from released substrate must always be less than the amount of biomass lost through decay, results in the net loss of biomass associated with decay (Grady et al., 1999). The 4.57 term in the stoichiometric coefficient for aerobic growth of autotrophic bacteria is the theoretical oxygen demand associated with the oxidation of ammonia nitrogen to nitrate nitrogen (Grady et al., 1999).

The other electron acceptor included in the model is nitrate nitrogen, S_{NO} , which is represented in column $i = 9$. It is produced by aerobic growth of autotrophic bacteria and removed by heterotrophic growth under anoxic conditions. Although nitrite nitrogen is an intermediate formed during nitrification, for simplicity in modelling ASM1_{SP} assumes that nitrate is the only oxidized form of nitrogen present. The factor 2.86 in the

stoichiometric coefficient for denitrification is the oxygen equivalence for conversion of nitrate nitrogen to nitrogen gas (N_2) and is included to maintain consistent units (Grady et al., 1999). Although not expressed explicitly in the model, nitrate nitrogen will also be removed by biomass decay which is accomplished by the recycling of organic matter during decay, making it available for anoxic growth of the heterotrophs (Henze et al., 1987).

Column $i = 10$ contains soluble ammonia nitrogen, S_{NH} , which is assumed to be the sum of the ionized (ammonium) and un-ionized (ammonia) forms. Ammonia nitrogen is utilized by growth of the biomass and is assumed to be released by biomass decay. The major sink for the ammonia nitrogen is as the energy source for the aerobic growth of autotrophic bacteria ($-1/Y_A$) (Henze et al., 1987). However, ammonia nitrogen is also incorporated into biomass during cell synthesis which is represented by the term ($-i_{N/XB}$) in the stoichiometric coefficient for S_{NH} in the growth process of both heterotrophic and autotrophic bacteria.

5.3.3 Processes in ASM1_{SP}

Listed in the leftmost column of Table 5.4 are the biological processes incorporated into the model, while their rate expressions are listed in the rightmost column. Three fundamental processes are considered: growth of biomass, decay of biomass and hydrolysis of particulate organics which are entrapped in the biofloc.

The first row in Table 5.4 represents the growth of heterotrophic bacteria under aerobic conditions. Soluble substrate is consumed which results in the production of heterotrophic biomass. Associated with this is the utilization of oxygen. Since COD units are used for both substrate and biomass and because oxygen can be considered as negative COD, continuity requires that the oxygen requirement equal the net COD removal (soluble substrate removed minus cells formed) (Henze et al., 1987). It is clear from the first row that ammonia nitrogen is utilized and incorporated into cell mass. The mathematical expression for the rate of the aerobic growth of the heterotrophs utilizes a

double nutrient limitation, with the concentrations of both readily biodegradable substrate and DO being rate determining, and a Monod-type saturation function is used to model the effect of each constituent. The primary purpose of the oxygen term in the rate equation is as a switching function which stops aerobic growth at low DO concentrations and thus the value of the saturation coefficient, $K_{O,H}$, is small (Henze et al., 1987). Similar to ASM1, ASM1_{SP} does not model directly the phenomenon of soluble substrate storage as removal of soluble substrate is considered to be proportional to growth. However, this event is handled in the model through the immediate entrapment of slowly biodegradable substrate.

The growth of heterotrophic bacteria under anoxic conditions (denitrification) is represented in row $j=2$ in Table 5.4. The process occurs at the expense of readily biodegradable substrate and results in heterotrophic biomass. Nitrate serves as the terminal electron acceptor and its removal is in proportion to the amount of readily biodegradable substrate removed minus the quantity of cells formed (Henze et al., 1987). Ammonia is removed from the solution and incorporated into the cell mass. Examining the rate equation for the denitrification process reveals that the effect of readily biodegradable substrate on the process rate is identical to the one for the aerobic growth represented in row $j=1$, including the value for the saturation coefficient, K_S . It is known that the maximum rate of substrate removal under anoxic conditions is often less than the one under aerobic conditions, which could either be because $\hat{\mu}_H$ is lower under anoxic conditions or because only a fraction of the heterotrophic biomass is able to denitrify (Henze et al., 2000). In order to handle this phenomenon, an empirical coefficient, η_g , where $\eta_g < 1$, is used in the rate equation for the denitrification process (Henze et al., 1987). Denitrification depends on the concentration of nitrate nitrogen in a manner analogous to the way in which aerobic growth depends on the dissolved oxygen concentration (Henze et al., 1987). Therefore, a Monod-type term is included in the rate equation to model the effect of nitrate concentrations. Because of the fact that anoxic growth is inhibited by the presence of dissolved oxygen, an inhibition term $\frac{K_{O,H}}{K_{O,H} + S_O}$ is

included in the rate equation. The coefficient $K_{O,H}$ has the same value as in the expression for aerobic growth so that as aerobic growth declines, anoxic growth increases (Henze et al., 1987).

Represented in row $j = 3$ in Table 5.4 is the aerobic growth of autotrophic biomass (nitrification). Ammonia nitrogen is removed from solution and used by the nitrifiers as the energy source for growth with the production of autotrophic cell mass and nitrate nitrogen as end products. In addition, a small amount of ammonia nitrogen is incorporated into cell mass and is represented by the term $-i_{N/XB}$ in the stoichiometric coefficient for S_{NH} . In order to model the dependency of the autotrophic specific growth rate upon the soluble concentrations of both ammonia nitrogen and oxygen, a double nutrient limitation using Monod-type saturation functions is used in the rate equation for nitrification.

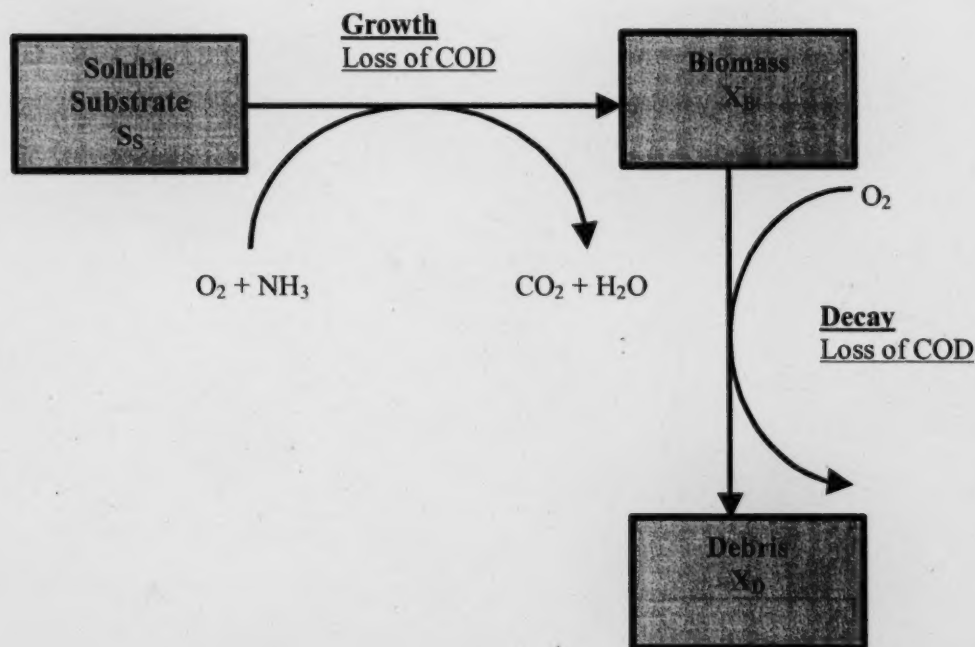


Figure 5.6 The traditional approach for modelling biomass decay (Adapted from Grady et al., 1999).

The most common technique to model biomass decay under aerobic conditions is using a rate expression which is first order with respect to the concentration of active biomass and to let the loss of COD due to biomass decay occurs at the expense of oxygen. A schematic diagram of this traditional approach is shown in Figure 5.6 in which both growth and decay of biomass are represented. Each unit of biomass COD lost results in the utilization of an equivalent amount of oxygen (Grady et al., 1999). Although this approach has worked well for the modelling of activated sludge systems performing only carbon oxidation and nitrification, it has been found that a modified approach would be warranted to model biomass decay under conditions in which oxygen is not the dominant electron acceptor (anoxic and anaerobic conditions) (Henze et al., 1987). For example, most studies suggest that biomass decay continues under anoxic conditions, at least for the fraction of the biomass that can denitrify (Grady et al., 1999).

In ASM1_{SP}, the approach adopted for modelling heterotrophic biomass decay is the death-regeneration concept of Dold et al. (1980), and is represented in row $j = 4$ in Table 5.4. A schematic representation of this approach is shown in Figure 5.7.

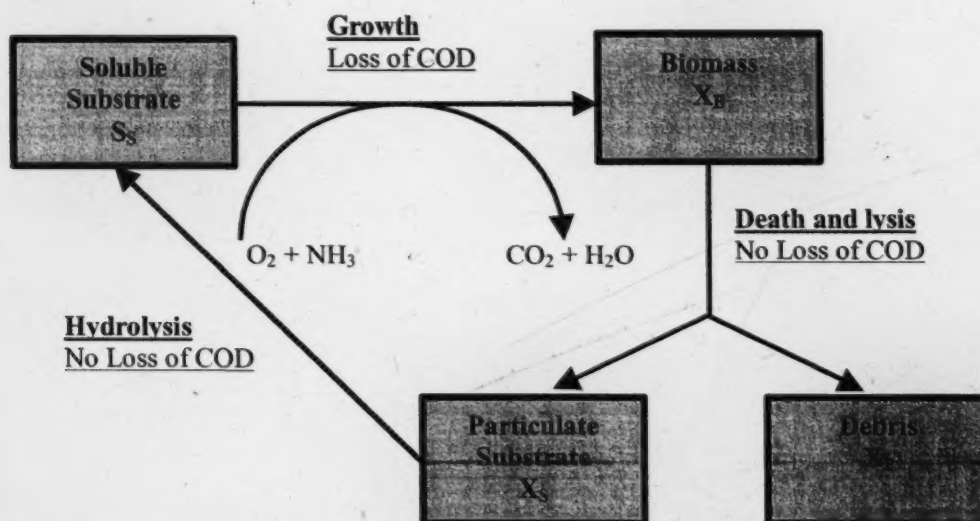


Figure 5.7 The death regeneration approach (Adapted from Grady et al., 1999).

It is clear from Figure 5.7 that decay acts to convert biomass to a combination of particulate products or biomass debris, X_p , and slowly biodegradable substrate and no loss of COD is involved in this transformation and no electron acceptor is utilized. According to this modelling approach, decay continues at a constant rate regardless of the environmental conditions (Henze et al., 1987). As depicted in Figure 5.7, the slowly biodegradable substrate produced from biomass decay is then hydrolysed, as represented in row $j = 6$ in Table 5.4, resulting in the release of an equivalent amount of readily biodegradable COD. In the case of aerobic conditions, that released substrate will be utilized to form new cells at the expense of dissolved oxygen, while in the case of anoxic conditions cell growth will occur with concomitant nitrate nitrogen uptake. Only under anaerobic conditions the released substrate will not be utilized and will accumulate.

Depicted in row $j = 5$ in Table 5.4 is the decay of autotrophic bacteria which is handled by $ASM1_{SP}$ in exactly the same manner as the decay of heterotrophs. The decay of the nitrifiers and the growth of the heterotrophs are interrelated as the released slowly biodegradable substrate produced from the autotrophic decay will be hydrolysed and then consumed by the heterotrophs.

As indicated by row $j = 4$ and row $j = 5$ in Table 5.4, $ASM1_{SP}$ assumes that biomass decay results in the production of nitrogen in the form of soluble ammonia instead of particulate organic nitrogen as in the case of $ASM1$. Because of this simplification two processes were eliminated: hydrolysis of particulate organic nitrogen into the soluble form and the conversion of soluble organic nitrogen into ammonia nitrogen (ammonification). These processes are very hard to characterize and a lot of uncertainty exists regarding the values for the parameters describing these processes in the literature (Grady et al., 1999).

Although hydrolysis is one of the most difficult processes to model in an activated sludge system, simple reaction rate expressions have been suggested by several researchers in which hydrolysis is assumed to be a first order process with respect to the concentration of heterotrophic biomass (Grady et al., 1999). In $ASM1_{SP}$ hydrolysis is

modelled in a manner that is similar to that in ASM1 with one simplification that ASM1_{SP} assumes that the hydrolysis process occurs at the same rate regardless of the environmental conditions. The process is depicted in row $j = 6$ in Table 5.4. In the rate expression of the hydrolysis process k_h is the hydrolysis coefficient (hr^{-1}) and K_X is a half-saturation coefficient ($\text{mg particulate substrate COD/mg active biomass COD}$). An important characteristic of this expression is that even though the rate is first order with respect to the heterotrophic biomass concentration, it is controlled by the ratio of particulate substrate concentration to heterotrophic biomass concentration, rather than by the particulate substrate concentration alone (Grady et al., 1999). This is necessary because the reaction is thought to be surface-mediated, depending on the presence of extracellular enzymes whose quantity is thought to be proportional to the biomass concentration (Grady et al., 1999).

5.4 STEADY STATE SIMULATIONS

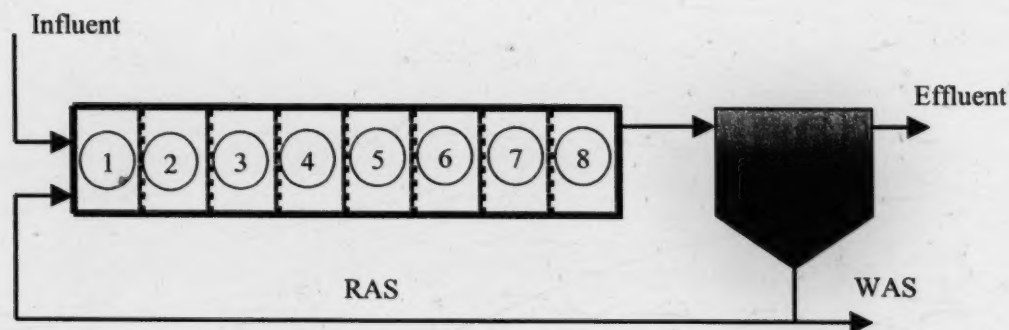


Figure 5.8 A schematic of the configuration of tanks used to represent the activated sludge process.

In order to gain insight into the process, steady-state simulations using ASM1_{SP} were conducted for activated sludge tanks #6, 7 and 8. Because of the unique configurations of this group of activated sludge tanks that were discussed earlier, the three tanks were modelled as one tank that has an aeration basin of a volume equivalent to the sum of the volumes of two aeration basin units at the GBWWTP and a secondary clarifier that has a volume equivalent to the sum of the volumes of three secondary

clarifier units. The aeration basin was modelled as eight continuous flow stirred tank reactors (CFSTRs) in series as depicted in Figure 5.8. Each of the eight CFSTRs has a volume of 2640 m³.

The steady state simulations were conducted for two operating modes: the first is plug flow and the second is step-feed mode. As shown in Figure 5.2, there are six feed points (represented by the thick white arrows) to each of the aeration tanks #6, 7 and 8: one at the entrance to the first pass; two in the second pass (start and middle) and three at the third pass (start, middle and end points). When any of activated sludge tanks #6, 7 and 8 is operated in the full step-feed mode, all the six gates are fully open. This configuration could not be simulated as flow measurements at these feed points were not available. Instead, an approximation had to be done in which only two equal feed streams were simulated: one to CFSTR #1 and the other to CFSTR #5.

Because the operating strategy at the GBWWTP is to control the dissolved oxygen concentration in the aeration basin at 2 mg/L, the dissolved oxygen concentration in the aeration basin was fixed at this value during the course of simulations. Because the operating strategy at the GBWWTP is to keep the return activated sludge (RAS) flow rate constant at 20 ML/d for each one of the activated sludge tanks at the plant, the steady state simulations for the group of tanks #6, 7 and 8 were conducted at a RAS flow rate of 60 ML/d.

Simulations were conducted at the following influent flow rates: 40, 80 and 120 ML/d. The solids residence time (SRT) used in the simulations varied from one day to eight days. Finally it should be mentioned that only the aeration basin was modelled in these simulations, i.e., the secondary clarifier was not modelled in the steady state simulations.

5.4.1 Characterization of the Wastewater

Listed in Table 5.8 are the characteristics of the influent wastewater which were used in the simulations. These characteristics are considered to be representative of the wastewater following primary sedimentation at the GBWWTP.

Table 5.8 Wastewater characterization of the primary effluent.

Soluble components	Concentration	Particulate components	Concentration
S_I	30 mg/L as COD	X_S	235 mg/L as COD
S_S	270 mg/L as COD	$X_{B,H}$	0 mg/L as COD
S_O	2 mg/L as COD	$X_{B,A}$	0 mg/L as COD
S_{NO}	0 mg/L as N	X_I	35 mg/L as COD
S_{NH}	36 mg/L as N	X_P	0 mg/L as COD

Listed in Table 5.6 are the values for the kinetic parameters in ASM1_{SP} that were used in the simulations while the values for the stoichiometric coefficients are listed in Table 5.7. These values are typical at neutral pH and 20°C for domestic wastewater (Grady et al., 1999).

5.4.2 Results – Plug Flow Mode

Presented in Figure 5.9 are the mixed liquor suspended solids (MLSS) concentrations in COD units in the eight CFSTRs representing the aeration basin for an influent wastewater flow of 80 ML/d. The MLSS concentration is the sum of the $X_{B,H}$, $X_{B,A}$, X_I and X_P concentrations. It is clear from Figure 5.9 that the MLSS profiles along the length of the aeration basin are constant. It is also evident that the MLSS concentration increases as the SRT increases. Presented in Figure 5.10 is the MLSS concentration in the last CFSRT (#8) at different SRTs and different influent flow rates.

It is clear from Figure 5.10 that the MLSS concentration increases as the influent wastewater flow increases.

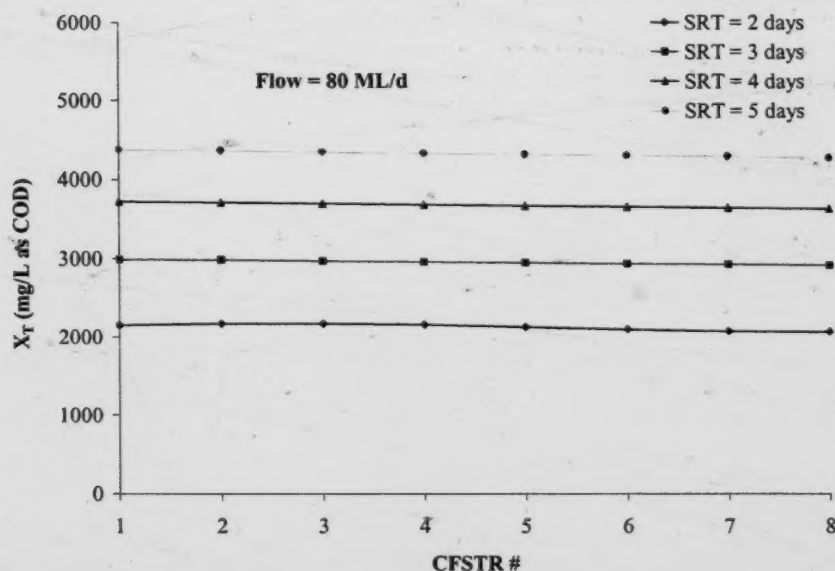


Figure 5.9 Mixed liquor suspended solids (MLSS) concentration profiles – Plug flow mode.

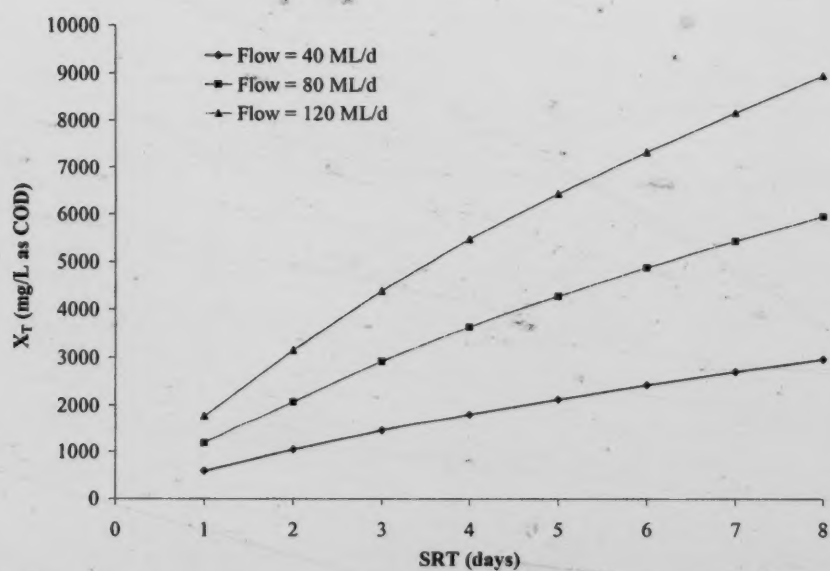


Figure 5.10 Mixed liquor suspended solids (MLSS) concentration in CFSTR #8 – Plug flow mode.

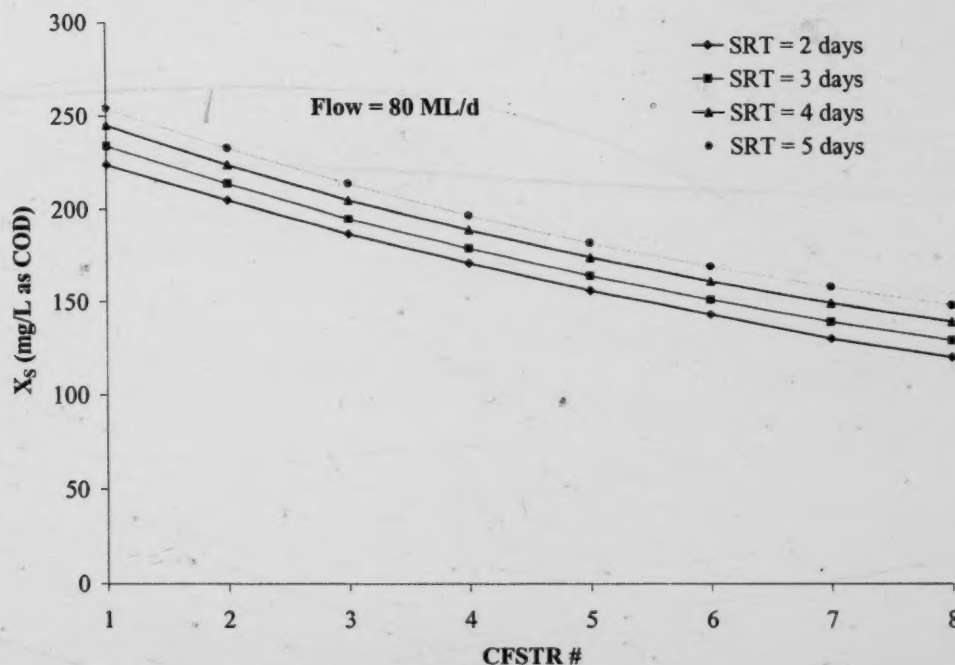


Figure 5.11 Slowly biodegradable substrate concentration profiles – Plug flow mode.

Shown in Figure 5.11 are the concentration profiles for the slowly biodegradable substrate at different SRTs for an influent wastewater flow of 80 ML/d. It is evident that the slowly biodegradable substrate has a declining concentration profile along the length of the aeration basin. However, because its contribution to the MLSS concentration is relatively small, the profiles shown in Figure 5.9 are still fairly constant. Figure 5.12 presents the readily biodegradable substrate (S_s) concentration in secondary effluent from which it is clear that almost complete utilization of this fraction of the substrate occurs at an SRT as low as one day. It is also clear from Figure 5.12 that increasing the SRT beyond two days does not have a pronounced effect on the S_s concentration in the secondary effluent.

The ammonia nitrogen concentration in secondary effluent is shown in Figure 5.13 while Figure 5.14 depicts the nitrate nitrogen concentration in secondary effluent for different SRTs.

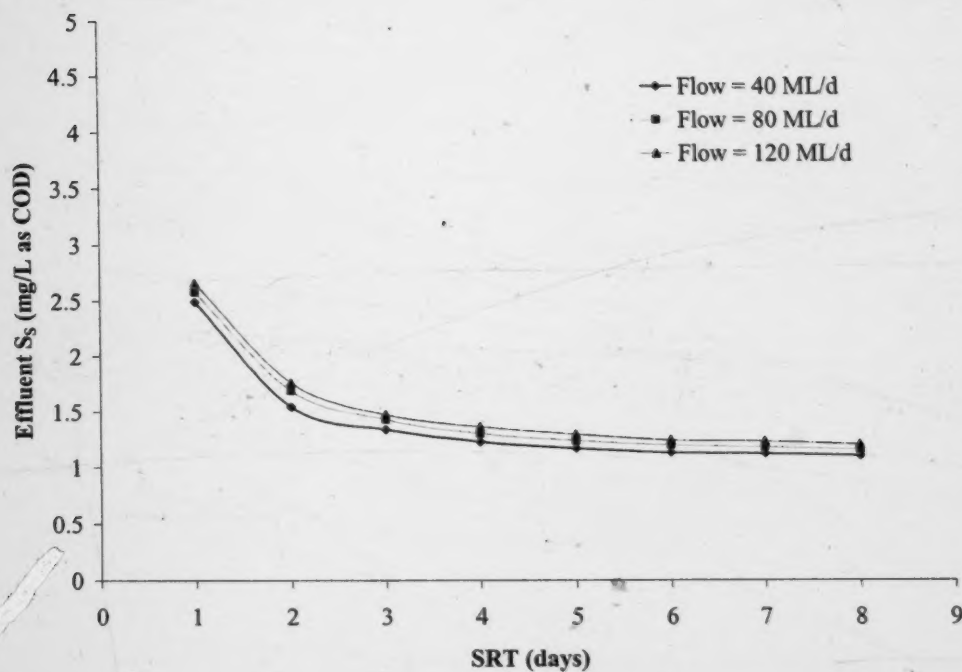


Figure 5.12 Readily biodegradable substrate in secondary effluent – Plug flow mode.

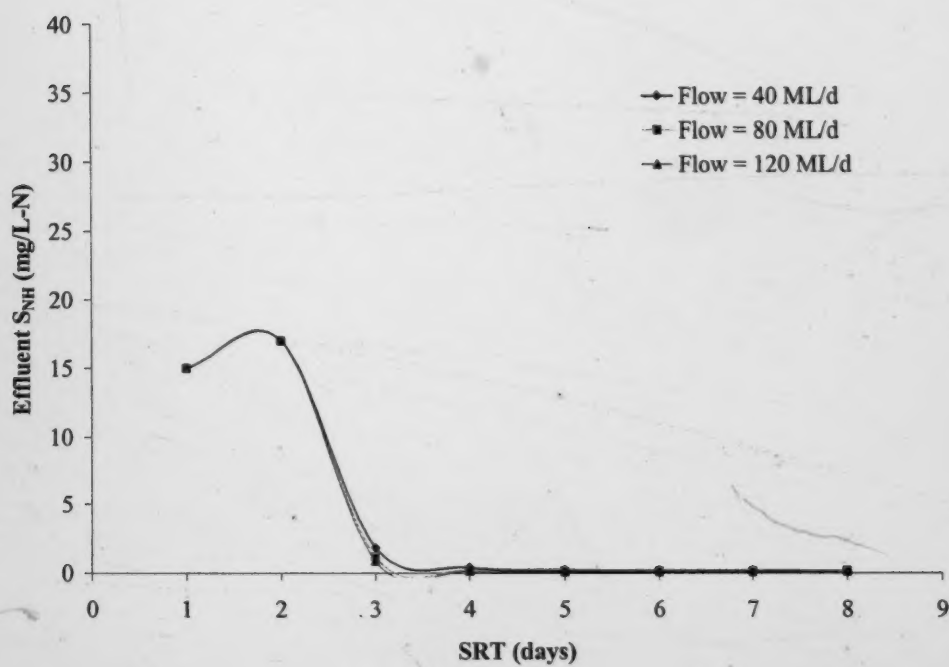


Figure 5.13 Ammonia concentration in secondary effluent – Plug flow mode.

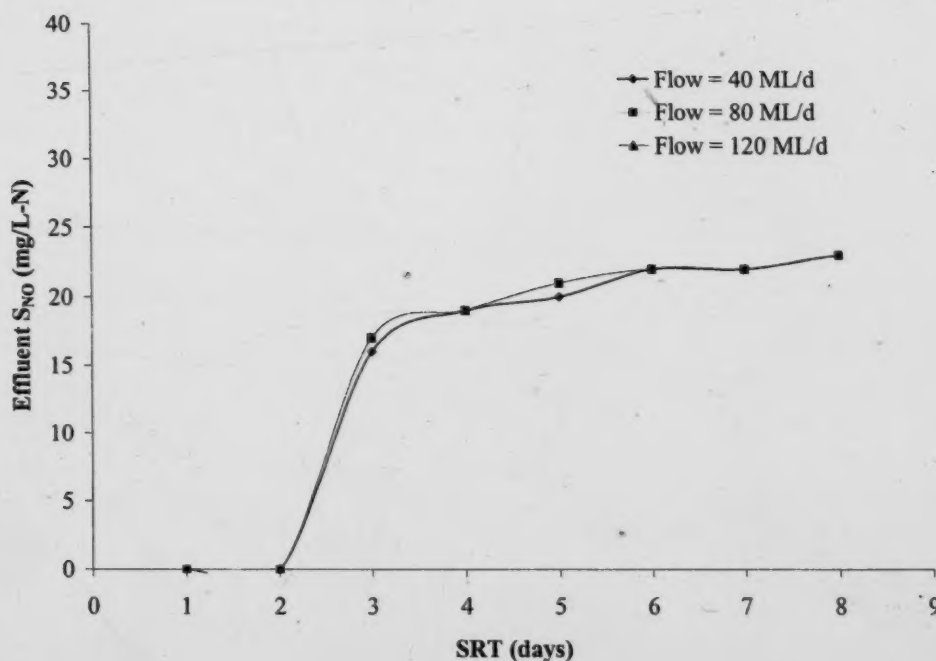


Figure 5.14 Nitrate concentration in secondary effluent – Plug flow mode.

The simulations shown clearly indicate that nitrification does not take place in the aeration basin when the SRT is less than two days. However, at an SRT of three days almost full nitrification can be established in the tank and increasing the SRT beyond three days does not have a pronounced effect on the percent nitrification. Comparison of the values used for μ for heterotrophic and autotrophic bacteria in Table 5.6 reveals that the value for autotrophs is almost an order of magnitude lower than that for heterotrophs, suggesting that the minimum SRT for nitrifying bacteria is almost an order of magnitude larger (Grady et al., 1999).

The value used for the half-saturation coefficient, which is the substrate concentration at which the bacteria grow at half of their maximal rate, for the autotrophs in Table 5.6 is 1.0 mg/L-N. This value is very low, thus nitrification will take place whenever the SRT is large enough to ensure stable growth and the ammonia-N concentration will rise rapidly as the SRT is decreased to the point of washout (Grady et

al., 1999). This explains the reputation of being all-or-none phenomenon that nitrification has gained (Grady et al., 1999). This phenomenon is clear in the simulations shown in Figure 5.13 and Figure 5.14.

5.4.3 Results – Step-Feed Mode

The mixed liquor suspended solids (MLSS) concentration profiles for an influent wastewater flow of 80 ML/d (40 ML/d to CFSTR #1 and 40 ML/d to CFSTR #5) are shown in Figure 5.15 from which it is clear that the MLSS profiles along the length of the aeration basin have a sharp drop between CFSTR #4 and 5. This is due to the fact that two equal influent flow streams were simulated, one fed to CFSTR #4 and the other fed to CFSTR #5. In the actual full step-feed configuration shown in Figure 5.2, there are six feed points. This actual configuration would have made the MLSS profiles decline gradually along the length of the aeration basin. However, the MLSS in the first and last CFSTR should not differ between the actual and simulated configurations.

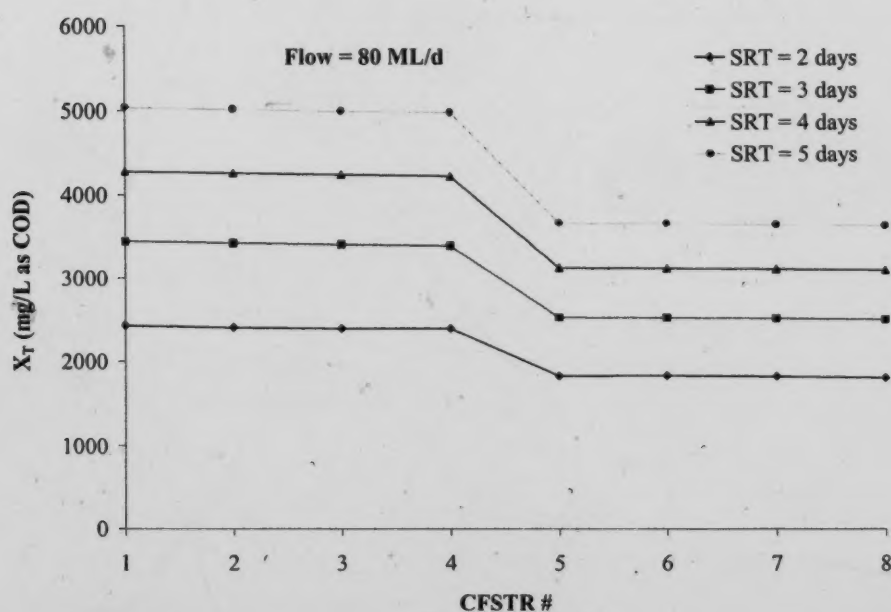


Figure 5.15 Mixed liquor suspended solids (MLSS) concentration profiles – Step-feed mode.

Figure 5.16 depicts the MLSS concentration in the last CFSRT (#8) at different SRTs and different influent flow rates from which it is clear that the MLSS concentration increases as the influent wastewater flow increases. Comparing Figure 5.16 to Figure 5.10 shows how the MLSS concentration leaving the aeration tank decreases when the mode of operation is changed from plug flow to step-feed.

The readily biodegradable substrate concentration in secondary effluent is depicted in Figure 5.17 from which it is clear that almost complete utilization of this fraction of the substrate occurs at an SRT as low as one day. Comparing Figure 5.17 to Figure 5.12 shows that only very slight increase in the readily biodegradable substrate concentration in the secondary effluent would occur if the mode of operation is switched from plug flow to step-feed.

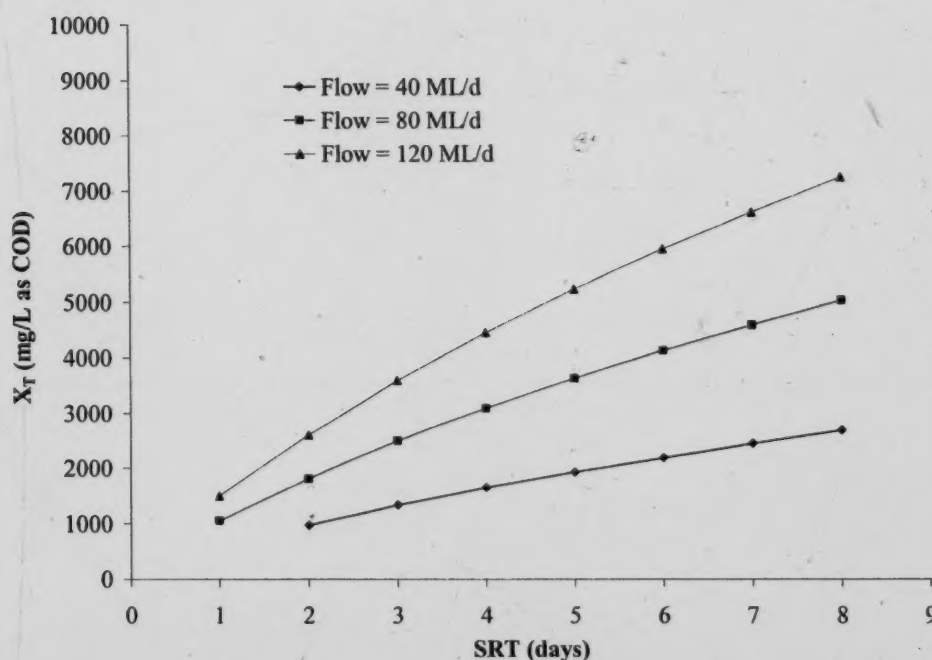


Figure 5.16 Mixed liquor suspended solids (MLSS) concentration in CFSTR #8 – Step-feed mode.

The ammonia nitrogen concentration in secondary effluent is shown in Figure 5.18 while Figure 5.19 depicts the nitrate nitrogen concentration in secondary effluent for different SRTs. Similar to the results of the steady state simulations for the plug flow mode of operation, the step-feed simulation results clearly indicate that nitrification does not take place in the aeration basin when the SRT is less than two days. At an SRT of three days almost full nitrification can be established in the tank and increasing the SRT beyond three days does not have a pronounced effect on the percent nitrification. Comparing Figure 5.18 to Figure 5.13 shows that, at an SRT of three days, there is a slight increase in the ammonia-N concentration in the secondary effluent when the step-feed mode of operation is used instead of the conventional plug flow, however, beyond an SRT of three days both Figure 5.18 and Figure 5.13 are virtually identical.

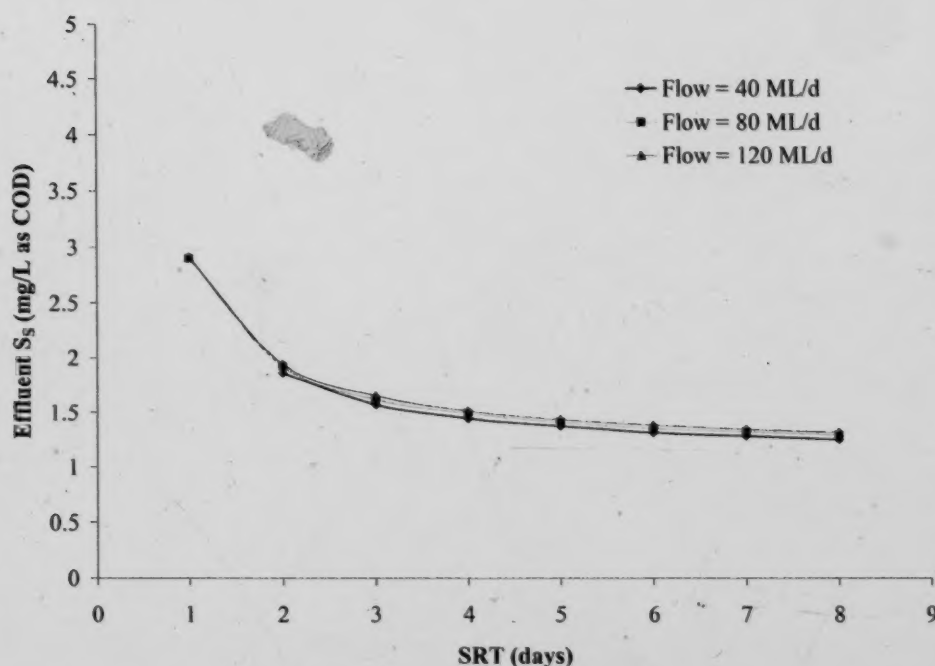


Figure 5.17 Readily biodegradable substrate in secondary effluent – Step-feed mode.

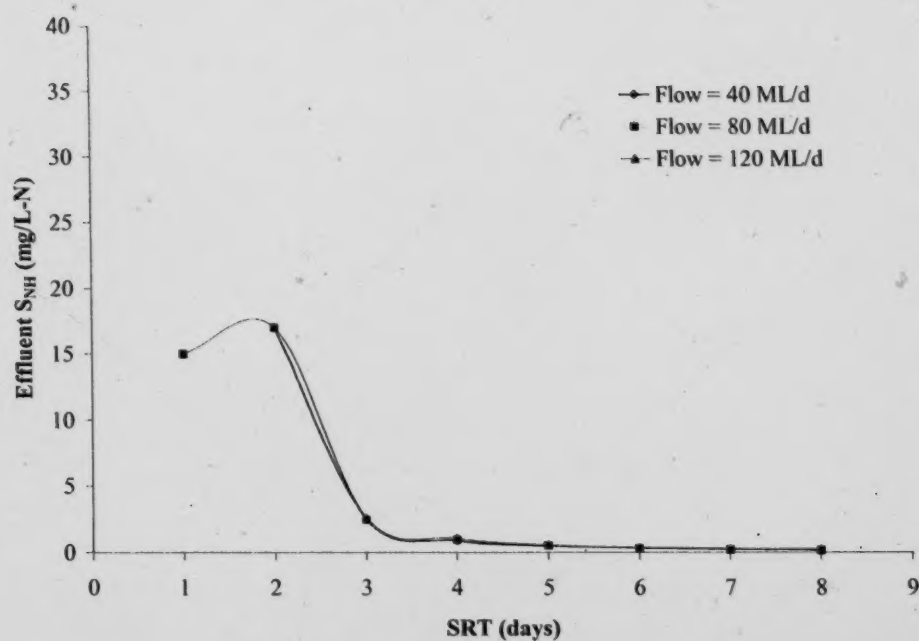


Figure 5.18 Ammonia concentration in secondary effluent – Step-feed mode.

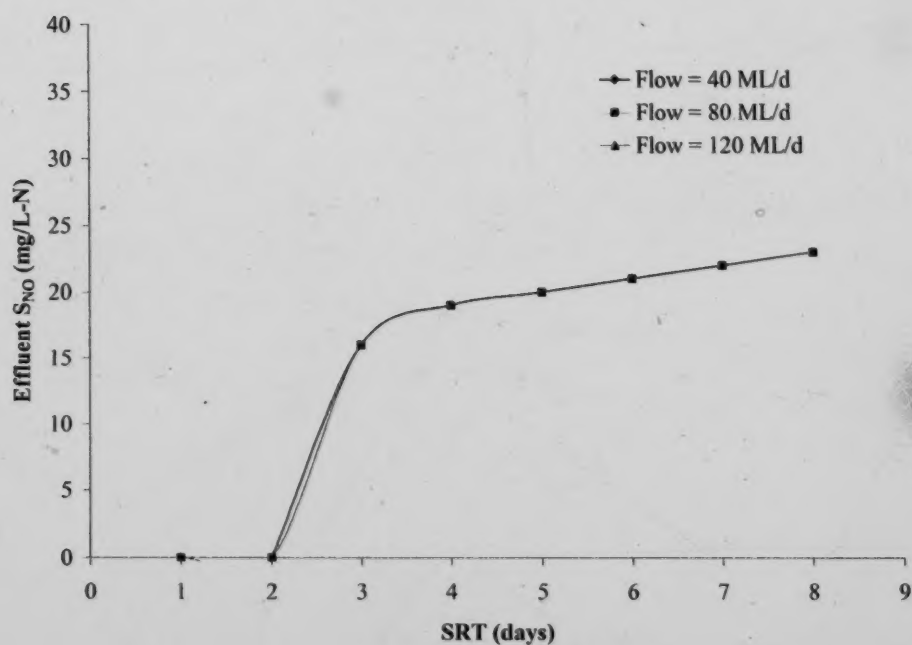


Figure 5.19 Nitrate concentration in secondary effluent – Step-feed mode.

5.5 DYNAMIC MODELLING UTILIZING THE EXPERIMENTAL DATA – RESULTS AND DISCUSSION

Dynamic modelling of the activated sludge tank was conducted using the experimental data that were collected during the two sampling surveys. Using the average values, listed in Table 5.2, for the quality parameters that were measured for the primary effluent, average concentrations were estimated in terms of the different components in ASM1_{SP}. This was done according to the procedure suggested by Grady et al. (1999) with two approximations. The first was that the soluble inert organic matter concentration in the primary effluent is equal to the average experimental value of the soluble COD in the secondary effluent minus the value of the soluble biodegradable substrate in the secondary effluent that was obtained from the steady state simulations at similar values of the solids residence time (SRT). This approximation assumes that all of the inert soluble substrate in the secondary effluent already existed in the primary effluent and passed the treatment without any change. In other words, no inert soluble organics originated from the decay of biomass. The second approximation assumed that soluble biodegradable organic nitrogen is readily available for the biomass in the form of ammonia nitrogen. This approximation was used in order to eliminate the need to simulate the ammonification process. The values of estimated primary effluent concentrations of the different components in ASM1_{SP} are listed in Table 5.9.

Typical diurnal patterns were drawn from the experimental data that were collected during the sampling surveys. These diurnal patterns were used as inputs to ASM1_{SP} and typical values for the stoichiometric coefficients listed in Table 5.7 were used. Values for the kinetic parameters were corrected for operating temperatures using the following equation:

$$r(t) = r(20) * \exp(\text{temp\_coeff} * (t - 20)) \quad [5.8]$$

where $r(t)$ is the value of the kinetic parameter at the operating temperature t , $r(20)$ is the base value of the kinetic parameter at 20°C and temp_coeff is the temperature coefficient

which was set equal to 0.069 (Grady et al., 1999). Typical values of the kinetic parameters that were listed in Table 5.6 were used as the base values at 20°C.

Table 5.9 Average primary effluent concentrations during the two sampling surveys.

Component	Concentration		Units
	First survey	Second survey	
S _I	33	45	mg/L as COD
S _S	241	175	mg/L as COD
S _O	2	2	mg/L as COD
S _{NO}	0.09	0.02	mg/L as N
S _{NH}	33.3	32.75	mg/L as N
X _S	208	155.5	mg/L as COD
X _{B,H}	0	0	mg/L as COD
X _{B,A}	0	0	mg/L as COD
X _I	40	31.5	mg/L as COD
X _P	0	0	mg/L as COD

Average values of the WAS flow, total suspended solids concentration in the RAS and total suspended solids concentration in the mixed liquor from the 4th pass of the aeration tank were used to calculate the solids residence time in the aeration tank. The SRT was calculated using the solids inventory in the aeration tank alone, i.e., it was assumed that no biological transformations occur in the secondary clarifier. Based on this assumption, the SRT calculated for the first and second sampling surveys were 2.3 and 2.1 days, respectively. A declining MLSS concentration profile similar to the one found from the steady state simulations for the step feed case was used to estimate the average MLSS concentration in the aeration tank based on the total suspended solids concentration in the mixed liquor collected from the 4th pass of the aeration tank.

As in the steady state simulations, the aeration tank basin was modelled as eight CFSTRs in series. The secondary clarifier was modelled as two compartments of equal

volume. These compartments are thought to model the clear water zone in the upper part of the secondary clarifier and they contain only soluble species and no biological processes are active in them (Henze et al., 1987). Under unsteady state conditions the volume of these compartments may dampen the peak concentrations in the effluent from the biological reactors. Based on the average values measured for the depth of the sludge blanket in the secondary clarifier during both surveys, approximately two thirds of the volume of the secondary clarifier can be assumed to be clean water while the other third is the depth of the sludge blanket in the tank. Because the operating strategy at the GBWWTP is to control the dissolved oxygen concentration in the aeration basin at 2 mg/L, the dissolved oxygen concentration in the aeration basin was fixed at this value during the course of the dynamic modelling.

Figure 5.20 shows the ammonia nitrogen concentration in the secondary effluent samples collected during the first survey. The predictions by ASM1_{SP} were in relatively good agreement with the experimental data. One discrepancy existed between the experimental data collected during the second survey and the results from the unsteady state simulations by ASM1_{SP}. This discrepancy is clear from Figure 5.21. The squares represent the nitrate nitrogen concentration in the secondary effluent samples collected during the second survey without modelling the denitrification process in the secondary clarifier. The predictions by ASM1_{SP} were consistently higher than the experimental data.

On the other hand, the actual ammonia nitrogen concentrations in the secondary effluent samples that were collected during the second survey were in relatively good agreement with the predictions by ASM1_{SP}. These findings suggested that denitrification was taking place in the secondary clarifier during the second survey. This conclusion was also supported by the fact that during the settling tests that were conducted by the plant operators on the mixed liquor samples from the 4th pass, the sludge had always risen after approximately 30 minutes from the beginning of a test, which was attributed to the release of nitrogen gas. Therefore, in the ASM1_{SP} modelling that was conducted on the experimental data that were collected during the second survey, it was decided to include

the denitrification occurring in the secondary clarifier as part of the activated sludge process configuration presented to ASM1_{SP}.

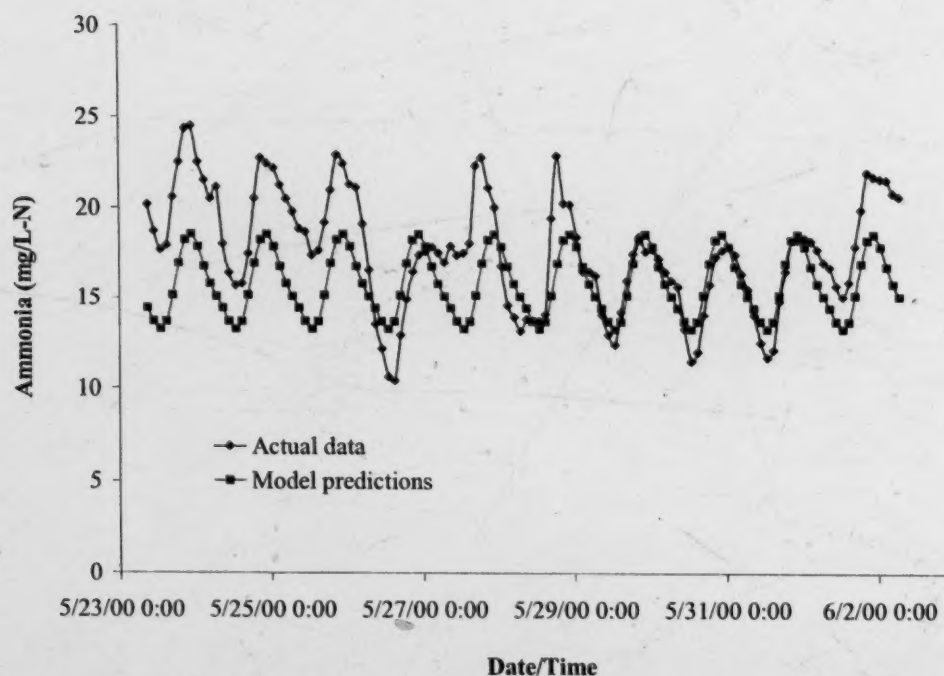


Figure 5.20 Ammonia nitrogen concentration in the secondary effluent – First survey.

The biological processes in the sludge blanket in the secondary clarifier were modelled by choosing the first biological reactor in the configuration file with the volume of the sludge blanket and the influent wastewater to the aeration tank was fed to the second bioreactor. The dissolved oxygen concentration in the sludge blanket in the secondary clarifier was assumed to be zero. As shown in Figure 5.21, after including the biological processes occurring in the secondary clarifier as part of the configuration, the predictions by ASM1_{SP} for the nitrate nitrogen concentration in the secondary effluent were in much better agreement with the experimental data.

Comparing Figure 5.22 to Figure 5.20 reveals that the agreement between the experimental data for the ammonia nitrogen concentration in the secondary effluent samples and the dynamic simulation by ASM1_{SP} for the second sampling survey was

relatively better than that found for the first survey. This could be partially due to the two rainfall events that were encountered during the 5th and 6th days of the first sampling survey. Another reason that accounted for the departure of the simulated concentration profile from the actual data collected during the first few days of the first sampling survey was a sludge bulking problem that took place in the secondary clarifier. The results of the activated sludge analysis that was performed by the laboratory of the GBWWTP indicated high filamentous length. The sludge bulking in the secondary clarifiers caused some of the microorganisms to escape with the effluent wastewater, which caused the quality of the effluent to deteriorate. In order to remedy this bulking condition, the staff in charge at the GBWWTP decided to seed the aeration tanks #6, 7 and 8 using activated sludge from the 4th pass of aeration tanks #9 and 10 which are biological nutrient removal (BNR) tanks. The seeding was conducted twice and in each time lasted for two hours. The objective was to enhance the settling characteristics of the activated sludge in the secondary clarifiers #6, 7 and 8.

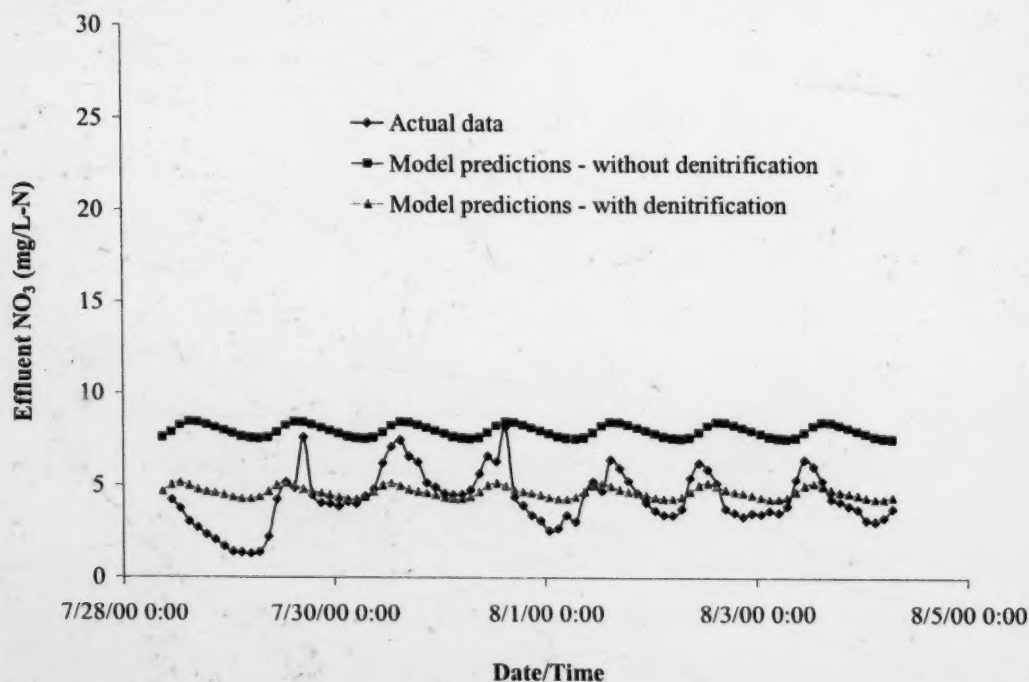


Figure 5.21 Nitrate nitrogen concentration in the secondary effluent – Second survey.

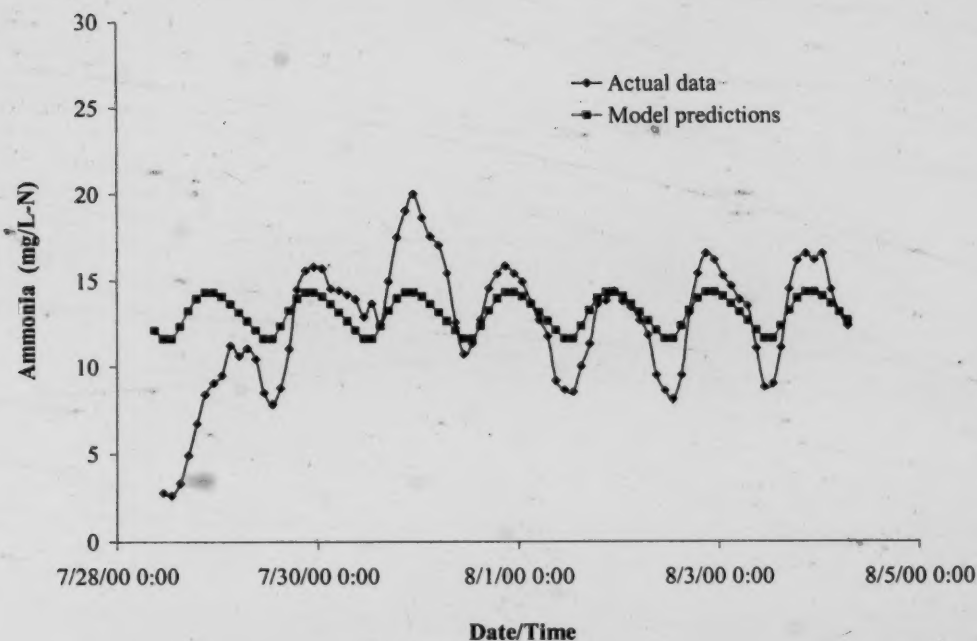
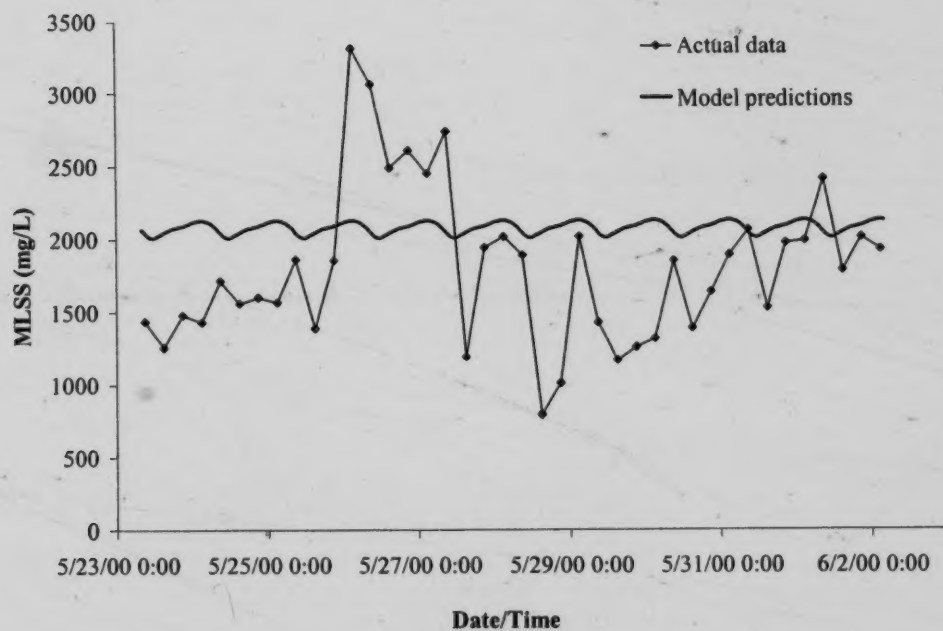
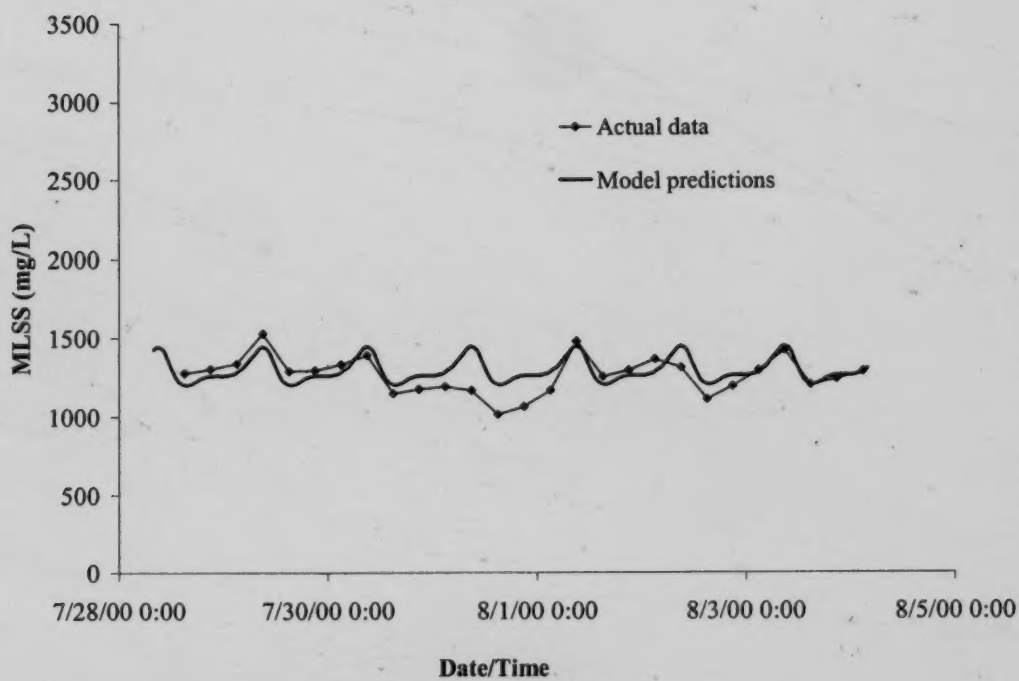


Figure 5.22 Ammonia nitrogen concentration in the secondary effluent – Second survey.

The suspended solids concentration in the mixed liquor samples that were collected from the 4th pass of aeration tank are shown in Figure 5.23 compared against the predictions by ASM1_{SP}. A conversion factor, $i_{O/XB,T} = 1.2 \text{ mg COD/mg TSS}$, was used to convert the MLSS predictions by ASM1_{SP} from the COD units into TSS units. This conversion factor is suggested by Grady et al. (1999) for converting the heterotrophic yield coefficient, Y_H , from the conventional mg biomass TSS/mg substrate COD removed units into the mg biomass COD/mg substrate COD removed units which are compatible with ASM1. It is clear from Figure 5.23 that the agreement between the experimental data and the dynamic simulation by ASM1_{SP} for the second sampling survey was relatively better than that found for the first survey. Again, this could be attributed to the two rainfall events and the sludge bulking condition that were encountered during the first sampling survey. As shown in Figure 5.23, during the 3rd and 4th days of the first sampling survey the actual MLSS concentrations were higher than the model predictions. This could be due to the seeding that was performed on the aeration tanks using the heavier mixed liquor from the BNR tanks at the GBWWTP.



(a) First survey



(b) Second survey

Figure 5.23 Mixed liquor suspended solids concentration.

5.6 CONCLUSIONS

Although the activated sludge tank was operated at an estimated SRT value of 2.1 days during the second survey while the estimated SRT value during the first survey was 2.3 days, partial nitrification took place in the aeration tank during the second survey. This was mainly due to the higher temperatures encountered during the second survey. The temperature of the mixed liquor from the 4th pass averaged 21°C while only averaged 16.9°C during the first survey. Based on the steady state simulations that were conducted at 20°C, a SRT value between two and three days will cause partial nitrification to take place in the aeration tank.

Visual experimental observations during the settling tests conducted on the mixed liquor from the 4th pass of the aeration tank as well as the unsteady state simulations by ASM1_{SP} indicated that denitrification took place in the secondary clarifier during the second survey, and therefore, the biological transformations occurring in the sludge blanket of the secondary clarifier had to be included as part of the activated sludge tank configuration. Although denitrification took place in the sludge blanket of the secondary clarifier during the second survey, the effluent quality in terms of total suspended solids did not deteriorate. This was probably because nitrification did not occur at full extent in the aeration tank, and therefore, the nitrate feed to the sludge blanket in the secondary clarifier did not produce nitrogen gas at a rate that was enough to cause the sludge flocs to rise to the effluent launders.

The modelling that was conducted herein gave insight into the activated sludge process and helped identify the occurrence of operational problems such as denitrification in the secondary clarifier. In addition, this type of modelling can be beneficial in predicting the quality of the secondary effluent discharged to the receiving stream.

5.7 REFERENCES

1. APHA-AWWA-WEF, Standard Methods for the Examination of Water and Wastewater, 19th Ed. American Public Health Association, American Water Works Association, and Water Environment Federation (Washington, DC, USA: American Public Health Association, 1995).
2. Barker, P.S. and P.L. Dold, "General Model for Biological Nutrient Removal Activated-Sludge Systems: Model Presentation", Wat. Env. Res. 69: 969-984 (1997).
3. Daigger, G.T. and D. Nolasco, "Evaluation and Design of Full-Scale Wastewater Treatment Plants Using Biological Process Models", Wat. Sci. Tech. 31(2): 245-255 (1995).
4. Dold, P.L., G.A. Ekama, and G.v.R. Marais, "A General Model for the Activated Sludge Process", Progress in Wat. Tech. 12(6): 47-77 (1980).
5. Eckenfelder, W.W., M.C. Goronszy, and T.P. Quirk, "The Activated Sludge Process: State of the Art", Critical Reviews in Environmental Control 15(2): 111-178 (1985).
6. Ganczarczyk, J.J., *Activated Sludge Process. Pollution Engineering and Technology No. 23* (NY, USA: Marcel Dekker, Inc., 1983).
7. Grady Jr., C.P.L., G.T. Daigger, and H.C. Lim, *Biological Wastewater Treatment*, 2nd ed. (NY, USA: Marcel Dekker, Inc., 1999).
8. Henze, M., C.P.L. Grady Jr., W. Gujer, G.v.R. Marais, and T. Matsuo, *Activated Sludge Model No. 1. IAWPRC Scientific and Technical Reports No. 1* (London, UK: IAWPRC, 1987).
9. Henze, M., W. Gujer, T. Mino, T. Matsuo, M.C. Wentzel, and G.v.R. Marais, *Activated Sludge Model No. 2. IAWQ Scientific and Technical Reports No. 3* (London, UK: IAWQ, 1995).
10. Henze, M., W. Gujer, T. Mino, and M. van Loosdrecht, *Activated Sludge Models ASM1, ASM2, ASM2d and ASM3. IAWQ Scientific and Technical Reports No. 9* (London, UK: IWA, 2000).

11. Lawrence, A.W. and P.L. McCarthy, "Unified Basis for Biological Treatment Design and Operation", J. San. Eng., ASCE 96(SA3): 757-778 (1970).
12. McKinney, R.E., "Mathematics of Complete-Mixing Activated Sludge", J. San Eng., ASCE 88(SA3): 87-113 (1962).
13. Mckinney, R.E. and R.J. Ooten, "Concepts of Complete Mixing Activated Sludge", *Trans. Of 19th Sanitary Eng. Conf.*, University of Kansas, USA, 1969.
14. Monod, J., "The Growth of Bacterial Culture", *Annual Review of Microbiology* 3: 371-394 (1949).
15. Zhao, H., O.J. Hao, T.J. McAvoy, and C.H. Chang, "Modelling Nutrient Dynamics in Sequencing Batch Reactor", J. Env. Eng., ASCE 123(4): 311-319 (1997).

CHAPTER 6. A HYBRID MODEL FOR A FULL-SCALE ACTIVATED SLUDGE PROCESS THAT INCORPORATES A MECHANISTIC AND ANN/TIME SERIES COMPONENTS*

6.1 INTRODUCTION

Mathematical models for wastewater processes can take many forms. At one extreme, some models are highly mechanistic. These are most useful to the researcher seeking to understand the events occurring in a system. Generally such models are deterministic and incorporate direct links between inputs and outputs through ordinary and partial differential equations that seek to mimic reactions mechanisms. Such models are large in scope and require a detailed knowledge of the system and the evaluation of a large number of parameters. An example is the Activated Sludge Model No. 1 (ASM1) developed by the International Water Association (IWA; formerly the International Association on Water Quality IAWQ). ASM1 contains 5 stoichiometric parameters and 14 kinetic parameters, some of which are difficult to estimate. Also, some of the processes included in mechanistic models are theoretical in nature. For example, the fermentation and hydrolysis processes in the ASM1. Deriving rate equations for such processes is difficult, thus rendering the calibration of the model more difficult. In order to use a mechanistic model like ASM1 in practice, a detailed characterization of the wastewater as well as a thorough calibration of the numerous model parameters included are essential. In some cases, tedious bioassays are needed in order to estimate some of the parameters.

At the other extreme, some models are highly empirical and they constitute the so-called "black box" class of models. They can be stochastic and can reflect real world responses. Examples of the class of black box models are time series models and Artificial Neural Networks (ANNs). They are, however, highly system specific, and as a result, they have to be developed specifically for the set of data under study, and hence, they are not easily transportable to new situations. Although the identification of the model structure for such types of models could be in some cases a tedious job, estimating

A version of this chapter has been submitted for publication. Gamal El-Din, A. and D.W. Smith, Water Research (June 2002).

the model parameters can be easily done with the modern developments in computing technology. Regular updating of these models, i.e., re-estimating the model parameters, can be easily and automatically performed, thus rendering these models highly adaptive. The adaptiveness of the black-box class of models renders them very attractive in the field of on-line real-time process control and optimization. In the current climate of higher power costs and increasing stringent regulations, real-time control of wastewater treatment processes will be the coming step in the wastewater treatment system development.

Many alternatives lie between the two extremes of highly mechanistic models and highly empirical models. These alternatives are generally called "gray-box" class of models. The objective of this study was to explore a hybrid gray-box model that can be used to describe the activated sludge process at the Gold Bar Wastewater Treatment Plant (GBWWTP), the largest plant in the Edmonton area. The model consists of two components; one is mechanistic and the other is a black-box component. The mechanistic component of the hybrid model is a simplified version of the ASM1. The black-box component is an ANN in some cases while it is a time series model in other cases. In order to validate the model, its predictions were compared with actual field data that were collected during two field-sampling surveys. Discussion on the modelling of the activated sludge process, description of the study, model development, modelling results, possible applications of the model are presented in the following sections.

6.2 MODELLING THE ACTIVATED SLUDGE PROCESS

Modelling of the activated sludge process has passed through a sequence of events: first, the removal of organic matter only; second, nitrification; third, nitrogen removal by biological denitrification; and forth, phosphorus removal by biological treatment of wastewater. Monod (1949) described the different growth phases of a bacterial culture and introduced the famous Monod kinetics that describe the relation between the exponential growth rate of bacteria and the concentration of a limiting nutrient; an empirical relationship that was the origin of mathematical models of continuous growth

microbial systems such as the ones developed by McKinney (1962), Lawrence and McCarthy (1970), and Eckenfelder (1985).

Realizing the benefits to be drawn from mathematical modelling, the International Water Association formed a task group in 1983 to promote the development of practical models to the design and operation of biological wastewater treatment systems (Henze et al., 1987). Activated Sludge Model No. 1 (ASM1) was the model developed by the IWA task group in 1987 to predict the performance of an activated sludge system carrying out carbon oxidation, nitrification and denitrification (Henze et al., 1987). ASM1 has proved to be an excellent tool for modelling nitrification-denitrification processes (Barker and Dold, 1997, Daigger and Nolasco, 1995, Henze et al., 1995 and Zhao et al., 1997).

In order to model biological nutrient removal in complex activated sludge processes, the model structure requires a high dimension and the model possesses a large number of stoichiometric and kinetic parameters. As a result, the models have grown more complex over the years, from ASM1, including nitrogen removal processes, to ASM2, including biological phosphorus removal processes and to ASM2d including denitrifying PAOs. The ASM1 contains 8 processes and 12 soluble and particulate components while the full version of ASM2 contains 19 processes and 19 components.

Scientific research and model application in engineering practice have different goals. Whereas the detailed structure of the models is used in order to describe new mechanisms which have been identified in advanced research projects, manageable models with a moderate number of parameters but a high potential to predict system behavior should be the basis for model application in engineering practice (Henze et al., 2000). Realizing this fact, in 1998 the IWA task group decided to develop a new modelling platform, the Activated Sludge Model No. 3 (ASM3), in order to create a tool for use in the next generation of activated sludge models (Henze et al., 2000). The ASM3 was published in 2000 and was designed to satisfy primarily the requirements for model application. At the same time it still keeps as many details as are necessary to obtain some insight into the interconnected processes (Henze et al., 2000). ASM3 relates to the

same dominating phenomena as does ASM1: oxygen consumption, sludge production, nitrification and denitrification in activated sludge systems treating wastewater of primarily domestic origin. ASM3 is designed to be the core of many different models. The complexity of ASM3 is comparable to ASM1, however, there is a shift of emphasis from hydrolysis to storage of organic substrates, a process which has been postulated and observed by many researchers (Henze et al., 2000). In the current study a simplified version of ASM1 was used as the mechanistic component of the hybrid model. In the following sections, this model component will be denoted by "ASM1_{SP}".

6.3 STUDY

The GBWWTP was constructed in 1956 on the southwest shore of the North Saskatchewan River. The present capacity of the plant is 950 ML/d for primary treatment and 420 ML/d for secondary treatment based on both hydraulic and process capacities. The GBWWTP provides both primary and secondary treatment for the incoming raw sewage. Primary treatment consists of grit removal, mechanical screening, and primary sedimentation. The secondary treatment provides biological treatment in a suspended growth activated sludge system, final settling and microorganism reduction. Detailed description of the GBWWTP was provided in the previous chapters.

At the GBWWTP there are ten rectangular aeration basins which are numbered 1 through 10. Each aeration basin has a corresponding rectangular secondary clarifier. Listed in Table 5.1 are the physical dimensions of the activated sludge tanks. In the current study, activated sludge tank #8 was sampled. The tank is part of a unique configuration at the GBWWTP that is composed of activated sludge tanks #6, 7 and 8. This configuration is shown in Figure 5.1. The three tanks are identical in dimensions and configuration. The mixed liquor from the three aeration basins is combined in a distribution chamber that distributes the mixed liquor among the three secondary clarifiers #6, 7 and 8. The return activated sludge (RAS) from each of the three clarifiers is combined in a distribution chamber that divides the RAS among the three aeration tanks. The reason that aeration tank #7 is in dotted line format in Figure 5.1 is that this

bioreactor is normally out of service while all the three secondary clarifiers are normally in service. This was the case in both of the sampling surveys that were conducted. The operating strategy at the GBWWTP is to control the value of the mixed liquor suspended solids (MLSS) in the aeration tank by changing the flow rate of the waste activated sludge (WAS) and keeping the flow rate of the RAS from each of the secondary clarifiers #6, 7 and 8 at the same value of approximately 20 ML/d. By controlling the value of the MLSS in the aeration tank, the operator can control the solids residence time (SRT) of the activated sludge process. At the GBWWTP, aeration tanks can be operated in either the plug flow mode or the step feed mode of operation.

During the months of July and August of each year, the normal strategy at the GBWWTP is to switch aeration tanks from plug flow mode (the normal mode of operation) to step feed mode in order to accommodate the higher wastewater flow rates usually encountered during these two months. The white thick arrows in Figure 5.2 represent the points of feed of the primary effluent to the aeration basin. The operating strategy at the GBWWTP is to control the dissolved oxygen concentration in the aeration basin at 2 mg/L at all times. After mixing and aeration, mixed liquor is discharged into a short distribution channel which feeds the mixed liquor into the secondary clarifiers. At Gold Bar, rectangular secondary clarifiers are divided lengthwise into a number of cells. Each cell is equipped with a chain and flight collector and scum trough. Secondary Clarifiers #6, 7 and 8 each have four cells. Each cell has two manual inlet slide gates located at the south end for distributing mixed liquor into the tank. Just beyond the inlet slide gates is a wooden baffle wall. The baffle helps maintain quiescent conditions and reduces short-circuiting in the clarifiers. Final effluent from each secondary clarifier flows over effluent V-notch weirs into common channels which combine together to direct the wastewater flow to the Ultra violet (UV) building.

6.3.1 The Sampling Campaign

Samples were collected from both the primary effluent entering aeration tank #8 and the secondary effluent from secondary clarifier #8. Samples were also collected from

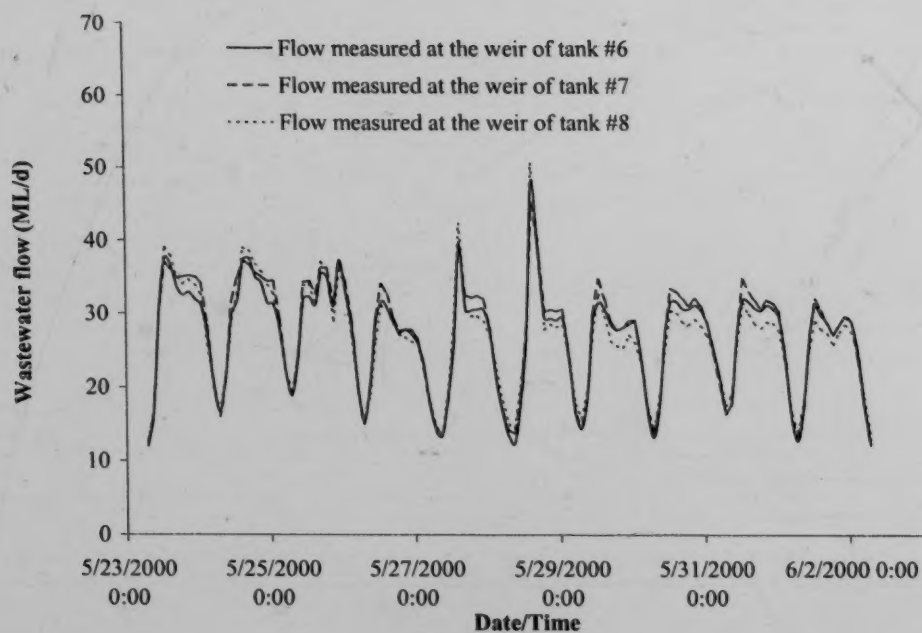
the mixed liquor in the 4th pass of the aeration tank and from the RAS line out of secondary clarifier #8. Two sampling surveys were conducted. The first survey started on May 23, 2001 at 9:00 hours and ended June 2, 2001 at 7:00 hours while the second survey started on July 28, 2001 at 9:00 hours and ended August 2, 2001 at 7:00 hours. Description of the sampling surveys was provided in detail in the previous chapter.

The first survey started on May 23, 2001 at 9:00 hours and ended June 2, 2001 at 7:00 hours. During this period all the activated sludge tanks at the plant were operated in the plug flow mode. During this survey the air temperature averaged 11.8°C while the temperature of the primary effluent averaged 16.6°C . Presented in Figure 6.1 (a) are the data for the wastewater flow measured at the effluent weir of each of the secondary clarifiers #6, 7 and 8 from which it is evident that the wastewater flow values measured for each of the three clarifiers were very close indicating that the flow was distributed approximately equally among the three tanks. The wastewater flow measured at the effluent weir of secondary clarifiers #6, 7 and 8 averaged 26.7, 27.4 and 26.6 ML/d, respectively. Two rainfall events took place during the 5th and 6th days of the survey, as shown in Figure 6.1 (a), during which the flow of wastewater measured at the effluent weirs of the clarifiers increased substantially above normal dry weather values. The TSS in the primary effluent averaged 95 mg/L while the total COD averaged 521 mg/L. Figure 6.2 (a) presents the COD data for the primary effluent samples collected from which it is evident that a weak diurnal pattern existed in the data. The soluble COD averaged 274 mg/L. The TKN and ammonia-nitrogen data for the primary effluent samples collected are shown in Figure 6.3 (a). The TKN in the primary effluent averaged 40.5 mg/L-N while the ammonia-nitrogen averaged 29.5 mg/L-N. It is clear from Figure 6.3 (a) that both the TKN and ammonia-nitrogen in the primary effluent followed the same general trend. A strong diurnal pattern existed in the data presented in Figure 6.3 (a).

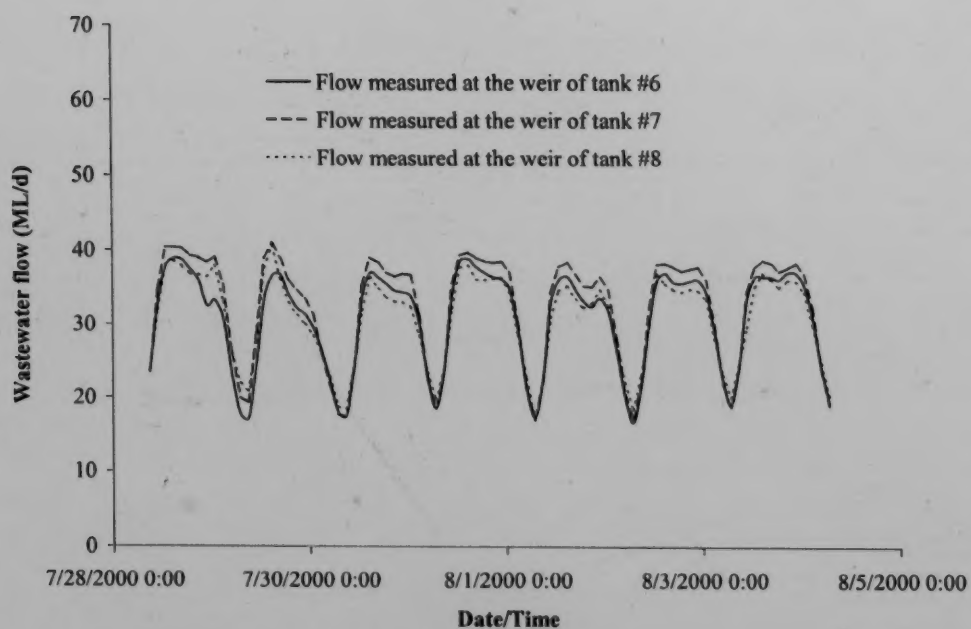
The second survey started on July 28, 2001 at 11:00 hours and ended August 4, 2001 at 7:00 hours. The step feed mode was used for operating aeration tanks #6, 7 and 8 during this survey. During this survey the air temperature averaged 20.9°C while the

temperature of the primary effluent averaged 20.5°C . These temperatures are higher than those recorded during the first sampling survey. The data for the wastewater flow measured at the effluent weir of each of the secondary clarifiers #6, 7 and 8 are presented in Figure 6.1 (b). As was the case in the first sampling survey, the wastewater flow data measured at the effluent weir of each of the secondary clarifiers #6, 7 and 8 were very close indicating equal flow distribution. The wastewater flow measured at the effluent weir of secondary clarifiers #6, 7 and 8 averaged 30.8, 32.4 and 30.8 ML/d, respectively. These average flow values are higher than that encountered during the first sampling survey. The TSS in the primary effluent averaged 74.7 mg/L. Presented in Figure 6.2 (b) are the COD data for the primary effluent samples collected. The total COD averaged 407 mg/L while the soluble COD averaged 220 mg/L. These values are lower than the average values for the first survey. Both the total and soluble COD in the primary effluent followed the same general trend. As was the case in the first survey, a weak diurnal pattern existed in the primary effluent COD data. Presented in Figure 6.3 (b) are the TKN and ammonia-nitrogen data for the primary effluent samples collected. The TKN in the primary effluent averaged 39.5 mg/L-N while the ammonia-nitrogen averaged 29 mg/L-N. These values are very close to the average values for the first survey.

Average values for the parameters measured during the two surveys are listed in Table 5.2 and Table 5.3. While the ammonia-N concentration in the secondary effluent averaged 17.6 mg/L during the first sampling survey, it only averaged 12.5 mg/L during the second survey. This was because partial nitrification took place during the second survey. Although nitrification in the group of activated sludge tanks #6, 7 and 8 was not one of the treatment objectives at GBWWTP, partial nitrification took place in the second survey due to the elevated water temperature of the mixed liquor in the aeration tank. It will be shown later that the modelling results clearly indicate that denitrification was taking place in the secondary clarifiers during the second survey. This phenomenon was also observed during the settling tests that were conducted by the plant operators for the mixed liquor from the 4th pass during the period of the second survey. After approximately 30 minutes from the beginning of a test, sludge started to rise which was attributed to the release of nitrogen gas.

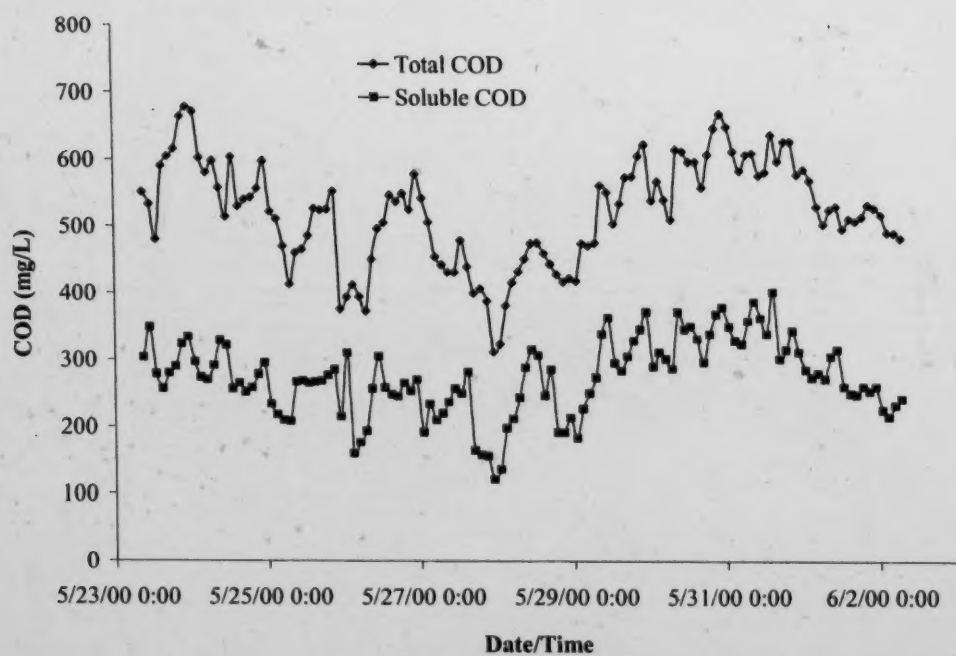


(a) First survey

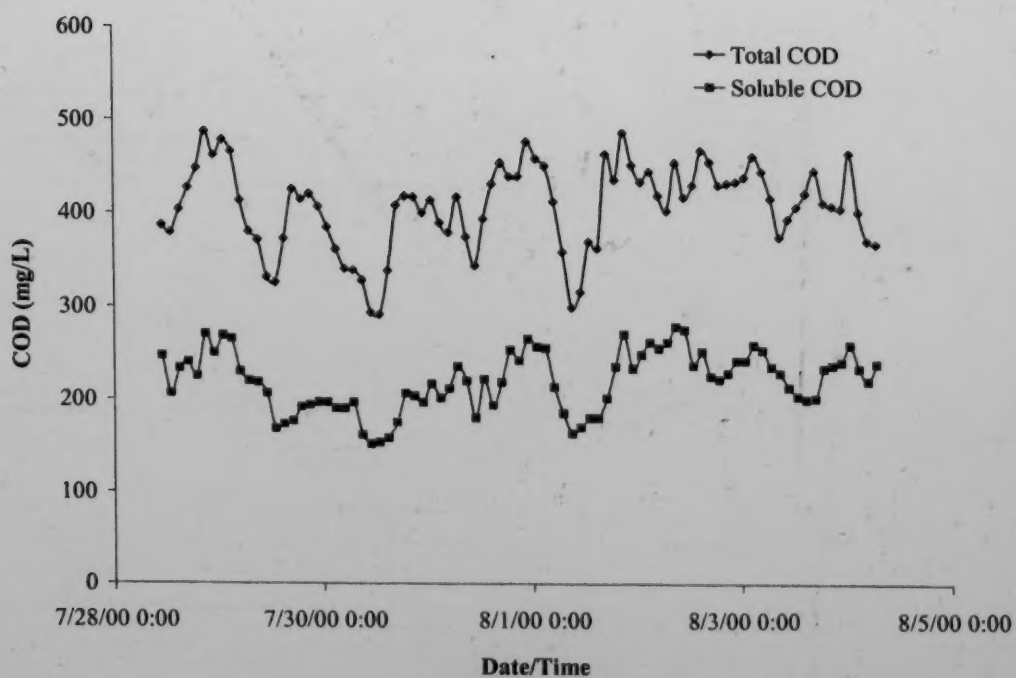


(b) Second survey

Figure 6.1 Wastewater flow measured at the effluent weir of each of the secondary clarifiers #6, 7 and 8.

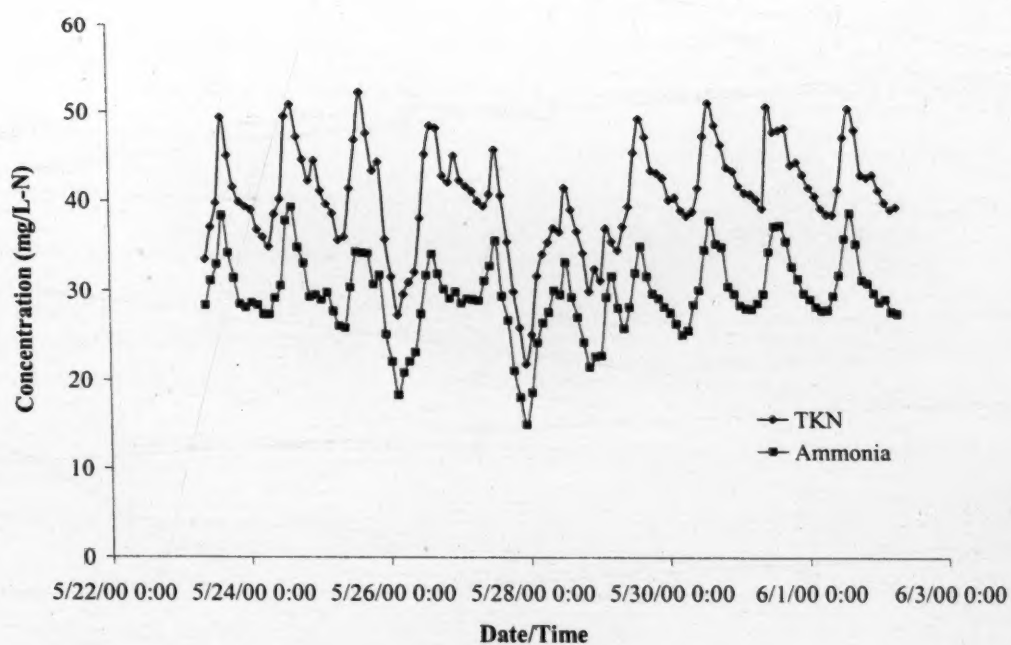


(a) First survey

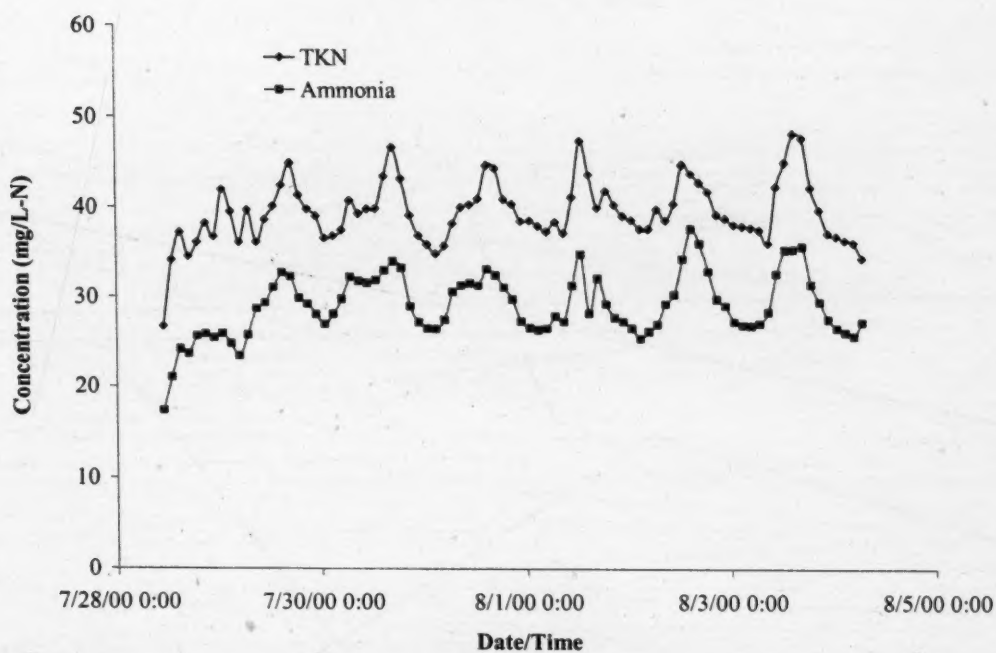


(b) Second survey

Figure 6.2 COD of the primary effluent samples.



(a) First survey



(b) Second survey

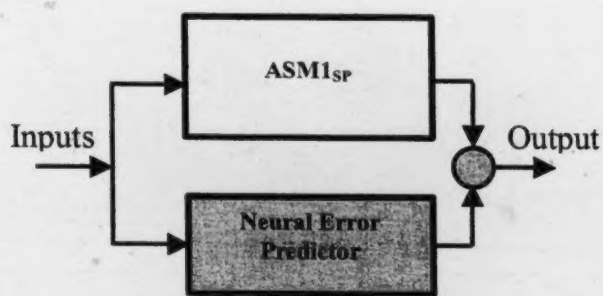
Figure 6.3 TKN and ammonia nitrogen concentrations in the primary effluent samples.

While the TSS concentration in the secondary effluent averaged 9.8 mg/L during the first sampling survey, it only averaged 2.4 mg/L during the second survey. During the second and third days of the ten-day period of the first sampling survey, the values of the secondary effluent TSS increased substantially above the average value. This was mainly due to a sludge bulking condition in the secondary clarifiers which was confirmed by the results of the activated sludge analysis performed by the laboratory of the GBWWTP which indicated high filamentous length. The sludge bulking in the secondary clarifiers caused some of the microorganisms to escape with the effluent wastewater, which caused the quality of the effluent to deteriorate. In order to remedy these bulking condition, the staff in charge at the GBWWTP decided to seed the aeration tanks #6, 7 and 8 using activated sludge from the 4th pass of aeration tanks #9 and 10 which are biological nutrient removal (BNR) tanks. The seeding was conducted twice and in each time lasted for two hours. The objective was to enhance the settling characteristics of the activated sludge in the secondary clarifiers #6, 7 and 8. Besides the change in the operational mode from plug flow in the first survey to step-feed in the second, the sludge bulking condition that occurred during the first survey and the partial nitrification-denitrification that took place during the second survey were the main differences between the conditions encountered during the first and the second sampling surveys that were conducted.

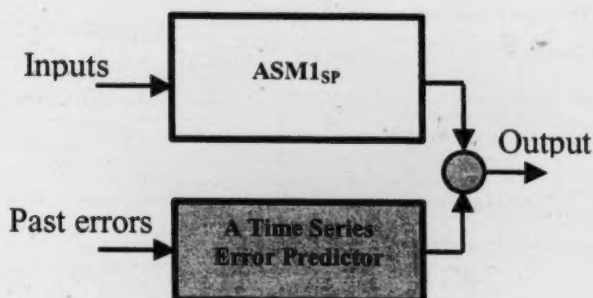
6.4 DESCRIPTION OF THE HYBRID MODEL

A schematic of the hybrid model is shown in Figure 6.4. The model consists of two components; a mechanistic component (a white-box component) and a black-box component. The mechanistic component "ASM1_{SP}" is a simplified version of the ASM1. ASM1_{SP} is a complex model, which has many kinetic and stoichiometric parameters that need to be estimated in order to calibrate the model. Some of these parameters are very difficult and time consuming to estimate. Therefore, it was decided to use typical values of the kinetic and stoichiometric parameters in ASM1_{SP} that are available in the literature for domestic wastewater after primary sedimentation and by incorporating a black-box component in the hybrid model, prediction errors can be compensated for. Therefore, the ASM1_{SP} provides only a preliminary prediction of the process behavior while the

function of the black-box component is to predict the error between the actual data and the predictions by $ASM1_{SP}$. As shown in Figure 6.4, two types of the hybrid model were used. The first uses a neural network model as the error-predictor while the second uses a time series model.



(a) $ASM1_{SP}$ + ANN component



(b) $ASM1_{SP}$ + Time series component

Figure 6.4 The hybrid model.

6.4.1 The Mechanistic Component of the Hybrid Model; $ASM1_{SP}$

$ASM1_{SP}$ has ten soluble and particulate components and six different processes. The model is presented in the matrix format in Table 5.4. Detailed description of the conceptual model, the different components and processes in the model, the values of the stoichiometric and kinetic parameters, and the simplifications in $ASM1_{SP}$ is provided in the previous chapter. ASIM (Activated Sludge SIMulation Program) was used in the

current study for solving the simultaneous mass balance equations for the constituents in ASM1_{SP}.

As clear from row $j = 4$ and row $j = 5$ in Table 5.4, ASM1_{SP} assumes that biomass decay results in the production of nitrogen in the form of soluble ammonia instead of particulate organic nitrogen as in the case of ASM1. Because of this simplification two processes were not simulated: hydrolysis of particulate organic nitrogen into the soluble form and the conversion of soluble organic nitrogen into ammonia nitrogen (ammonification). These processes are very hard to characterize and a lot of certainty exists regarding the values for the parameters describing these processes in the literature (Grady et al., 1999).

Although hydrolysis is one of the most difficult to model processes in an activated sludge system, simple reaction rate expressions have been suggested by several researchers in which hydrolysis is assumed to be a first order process with respect to the concentration of heterotrophic biomass (Grady et al., 1999). In ASM1_{SP} hydrolysis is modelled in a manner that is similar to that in ASM1 with one simplification that ASM1_{SP} assumes that the hydrolysis process occurs at the same rate regardless of the environmental conditions.

6.4.2 The Black-Box Component of the Hybrid Model

As was mentioned previously, two types of the hybrid model were utilized. The first uses an artificial neural network-type model as the black-box component in order to predict the error between the actual data and the predictions by the deterministic component ASM1_{SP}, while the second uses a Box-Jenkins time series-type model as the error-predictor component. Overview of ANN and the Box-Jenkins time series methodologies was provided in the previous chapters.

Among the many different architectures, the multi-layer perceptron architecture is commonly used for prediction, and therefore, was utilized in the current study. In the

current study, the input data were scaled linearly in the range $[-1,1]$ before being presented to the network. The logistic function was utilized as the transfer function in the hidden nodes. The linear function was used for transfer by the output nodes. In the current study, only one hidden layer was used and the number of hidden nodes was increased incrementally from 2 nodes to 10 nodes with increments size of one node and the parsimonious model that gave the best results was selected. In the modelling that was undertaken in this study, a batch-mode (weight updates were done after each epoch and not after each training pattern) back-propagation algorithm was used in the course of training and the network was saved at the point of minimum training error. Training was conducted using NeuroShell 2 software from Ward Systems Group Inc. In the present study training was stopped after 25, 50, 75, 100, 150, 200, 300, 400 and 500 epochs, for every candidate model that was tested. Each time training was stopped, the model was tested against the verification data set. It will be shown later that using this systematic approach, the point at which training should be stopped so that generalization could be maximized was identified. After training a neural network on the training data set, the network has to be validated or tested against another set of data that is usually called the "validation data set" that has never been seen before by the network during the course of training. The ability of the network to generalize from the training data set to the validation data set will determine its performance. In all the neural network modelling that has been conducted, the experimental data collected were split into two parts, one for training and the other for validation. After a model has been trained, the validation data set was used to judge the accuracy of the predictions generated by the trained network. This was done by calculating the R^2 value for the validation data set and comparing it to the value computed for the training data set. In each of the two sampling surveys conducted data for the last two days were used for validation while the rest of the data set was used for training.

The identification, estimation, and diagnostic checking of time series models were described in detail in Chapter 4. The same procedure of model building that was described in Chapter 4 was used herein. In the current study, a variety of checks were applied to each model, and the test results were considered as a group. The statistical

assumptions about the random error component a_t implied by the theoretical Box-Jenkins methodology are such that the model residuals should be white noise, in other words, should be uncorrelated and normally distributed around a zero mean. Residual diagnostics are tools by which these assumptions can be tested. Further more, models that have met these assumptions are compared using closeness-of-fit statistics applied to the residuals. Some of the statistics that were computed as part of the residual diagnostics are the residual mean (mean error) and mean percent error. Assuming that the form of the model is correct, the estimated autocorrelations of the residuals would be uncorrelated and distributed approximately normally about zero with variance $1/n$, and hence, with a standard error $1/\sqrt{n}$ (Box and Jenkins, 1976). Therefore, correlograms of the residuals (see Figure 6.5 for an example) were examined for correlations greater than two standard deviations since large correlations may have indicated model inadequacies, especially if they were at lower lags. The normality of residuals was examined by plotting the histogram and normal probability plots of the residuals. Among the closeness-of-fit statistics that were calculated are the mean absolute error, residual standard error, mean absolute percent error, and the index of determination (R^2). These are descriptive statistics that are useful for comparing different models that all passed the validation step. Besides examining the R^2 value, a time series plot of the measured output values and the values predicted by a model was also examined as an indication of the accuracy of the model. For each candidate model that has been tested, all of the above mentioned statistics were calculated. In addition, plots of the correlogram, histogram, and normal probability of residuals were drawn and white noise checks of the residuals were conducted in order to check the validity of the models. In all the time series modelling that has been conducted, time series data were split into two parts, one for estimating the model parameters (i.e. calibrating the model) and the other for validating (i.e. verifying) the model. After a model has been estimated, the validation data set was used to judge the accuracy of the forecasts generated by the estimated model. This was done by calculating the R^2 value for the validation data set and comparing it to the value computed for the estimation data set. In each of the two sampling surveys conducted, data for the last two days were used for validation while the rest of the data set was used for estimating the model parameters.

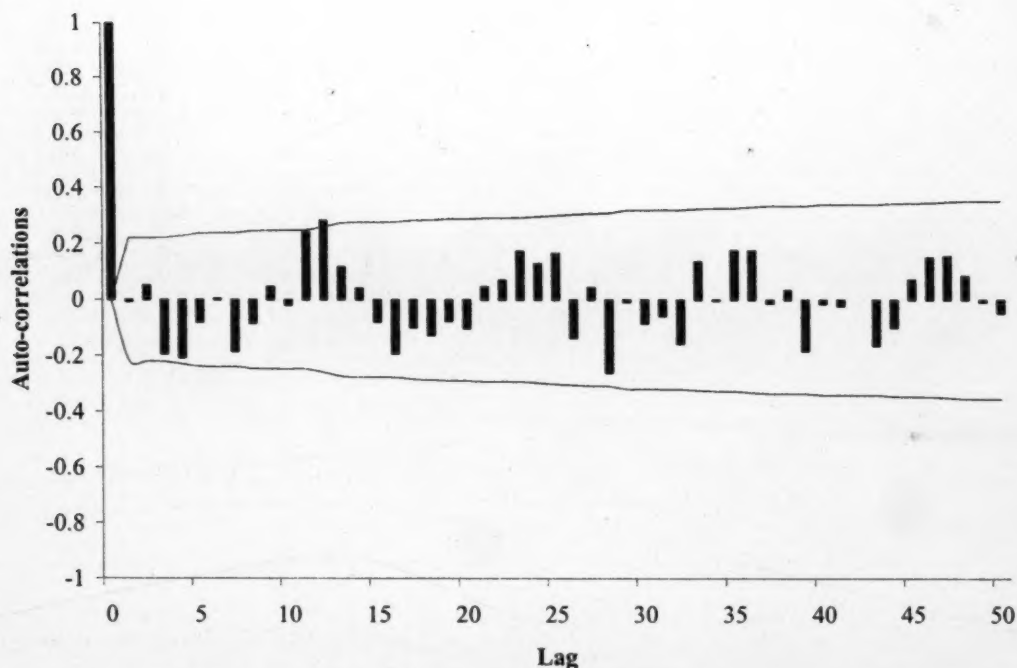


Figure 6.5 Correlogram of the residuals from the hybrid model that was used to predict the nitrate-nitrogen concentration in the secondary effluent for the second sampling survey. Solid lines represent the 95% confidence limits of two standard deviations.

6.5 MODEL DEVELOPMENT

As was mentioned previously, the hybrid models consisted of two components: a deterministic model component, $ASM1_{SP}$, which is a simplified version of Activated Sludge Model No. 1, and a black-box component or error-predictor component. The black-box component was an artificial neural network in some cases while it was a time series model in others.

6.5.1 $ASM1_{SP}$ - Unsteady State Simulations

The experimental data that were collected during the two sampling surveys were first compared to the predictions provided by $ASM1_{SP}$ alone (dynamic unsteady state

simulations) and then were compared to the predictions obtained by the hybrid model after including the black box component. Using the average values, listed in Table 5.2, for the quality parameters that were measured for the primary effluent, average concentrations were estimated in terms of the different components in ASM1_{sp}. This was done according to the procedure suggested by Grady et al. (1999) with two approximations. The first assumed that the soluble inert organic matter concentration in the primary effluent is equal to the average experimental value of the soluble COD in the secondary effluent minus the value of the soluble biodegradable substrate in the secondary effluent that was obtained from steady state simulations that were conducted using ASM1_{sp} at similar values of the solids residence time (SRT) using wastewater characteristics, which are considered to be representative of the wastewater following primary sedimentation at the GBWWTP, and typical values for the kinetic and stoichiometric parameters. According to this approximation, all of the inert soluble substrate in the secondary effluent already existed in the primary effluent and passed the treatment without any change. In other words, no inert soluble organics originated from the decay of biomass. The second approximation assumed that soluble biodegradable organic nitrogen is readily available for the biomass in the form of ammonia nitrogen. This approximation eliminated the need for the ammonification process. The values of estimated primary effluent concentrations of the different components in ASM1_{sp} are listed in Table 5.9.

Typical diurnal patterns were drawn from the experimental data that were collected during the sampling surveys. These diurnal patterns were used as inputs to ASM1_{sp} in order to perform the dynamic unsteady state simulations. Typical values for the stoichiometric coefficients listed in Table 5.7 were used. Values for the kinetic parameters were corrected for operating temperatures using the following equation:

$$r(t) = r(20) * \exp(\text{temp\_coeff} * (t - 20)) \quad [6.1]$$

where $r(t)$ is the value of the kinetic parameter at the operating temperature t , $r(20)$ is the base value of the kinetic parameter at 20°C and temp_coeff is the temperature coefficient

which was set equal to 0.069 (Grady et al., 1999). Typical values of the kinetic parameters that were listed in Table 5.6 were used as the base values at 20°C. Average values of the flow of waste activated sludge (WAS), total suspended solids concentration in the return activated sludge (RAS) and total suspended solids concentration in the mixed liquor from the 4th pass of the aeration tank were used to calculate the solids residence time (SRT) in the aeration tank. The SRT was calculated using the solids inventory in the aeration tank alone, i.e., it was assumed that no biological transformations occur in the secondary clarifier. Based on this assumption, the SRT calculated for the first and second sampling surveys were 2.3 and 2.1 days, respectively. Because of the unique configurations of activated sludge tanks #6, 7 and 8 that were discussed earlier, the three tanks were modelled as one equivalent tank that has an aeration basin of a volume equivalent to the sum of the volumes of two aeration basin units at the GBWWTP and a secondary clarifier that has a volume equivalent to the sum of the volumes of three secondary clarifier units. A schematic diagram of the configuration used to represent the equivalent activated sludge tank is shown in Figure 6.6 for a plug flow mode of operation.

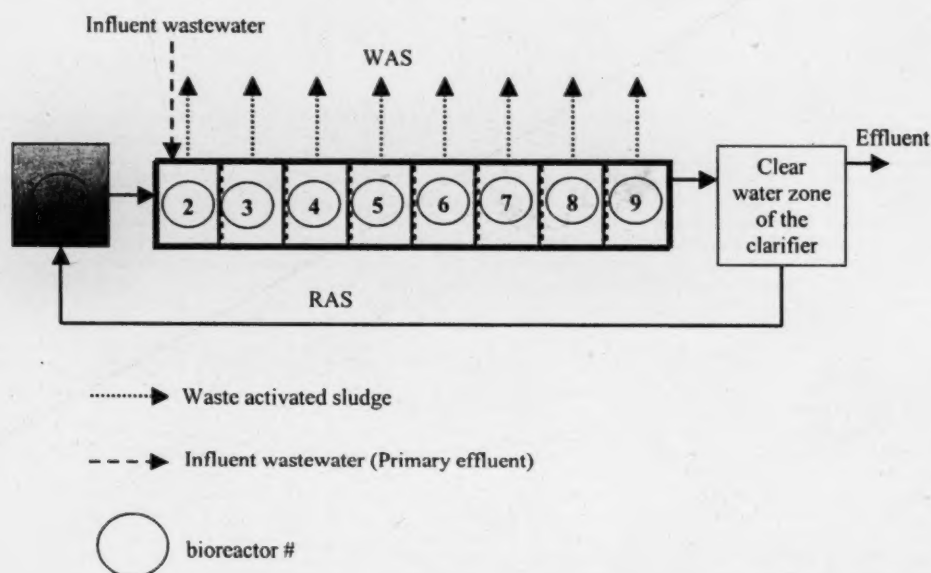


Figure 6.6 A schematic of the configuration used to represent the equivalent activated sludge tank - Plug flow mode of operation.

The aeration basin of the equivalent tank was modelled as eight continuous flow stirred tank reactors (CFSTRs) in series. The clear water zone of the secondary clarifier of the equivalent tank was modelled as one compartment. This compartment is thought to model the clear water zone in the upper part of the secondary clarifier and it contains only soluble species and no biological processes are active in it. Under unsteady state conditions the volume of this compartment may dampen the peak concentrations in the effluent from the biological reactors. Based on the average values measured for the depth of the sludge blanket in the secondary clarifier during both surveys, approximately two thirds of the volume of the secondary clarifier were assumed to be clear water while the other third represented the depth of the sludge blanket in the tank. Therefore, and because partial nitrification took place in the aeration tanks during the second sampling survey, the biological transformations in the sludge blanket had to be accounted for. In order to do so, a biological reactor that depicts the sludge blanket was added to the configuration file upstream of the eight CFSTRs representing the aeration basin of the equivalent activated sludge tank. This bioreactor is assigned #1 in Figure 6.6. The volume of this reactor was set equal to the volume of the sludge blanket of the secondary clarifier of the equivalent activated sludge tank and the dissolved oxygen concentration in this reactor was assumed to be zero. As indicated in Figure 6.6, the influent wastewater to the aeration tank was fed to CFSTR #2.

The unsteady state simulations for the first sampling survey was conducted using plug flow mode of operation while a step feed mode was used for the simulation conducted for the second sampling survey. As shown in Figure 5.2, there are six feed points (represented by the thick white arrows) to each of the aeration tanks #6, 7 and 8: one at the entrance to the first pass; two in the second pass (start and middle) and three at the third pass (start, middle and end points). When any of the activated sludge group of tanks #6, 7 and 8 is operated in the full step feed mode, all the six gates are fully open. This configuration could not be simulated as flow measurements at these feed points were not available. Instead, an approximation had to be done in which only two equal feed streams were simulated: one to CFSTR #2 and the other to CFSTR #6. Because the operating strategy at the GBWWTP is to control the dissolved oxygen concentration in

the aeration basin at 2 mg/L, the dissolved oxygen concentration in the aeration basin (CFSTRs #2-9) was fixed at this value during the course of the unsteady state simulations.

6.5.2 The Hybrid Model

A hybrid model that utilizes an artificial neural network as the error-predictor component was used to predict the ammonia nitrogen concentration in the secondary effluent. A schematic representation of the ANN error-predictor component is shown in Figure 6.7. The neural network takes the TKN, ammonia nitrogen, total COD, soluble COD and water temperature of the primary effluent (PE) at $t - 6$ and $t - 8$ hours as well as the last value measured for the mixed liquor volatile suspended solids in the 4th pass of the aeration tank as inputs in order to predict the ammonia nitrogen concentration in the secondary effluent at time t hours. For the first survey, the size of the hidden layer used was six neurons, while only three hidden neurons were needed in the neural network that was trained and tested against data collected from the second survey.

As was mentioned previously, the performance of the neural network was monitored during the course of training by stopping training and calculating the R^2 value for both the training and validation data sets. In this manner, the point at which training was stopped was visually determined from a graph such as that of Figure 6.8, which indicates that training was stopped after 100 epochs. Beyond this point, there was no further pronounced improvement in the network performance with respect to the training data set, however, the generalization ability of the network, measured by the value of the R^2 for the validation data set, started to decline.

A hybrid model that utilizes a time series model as the error-predictor component was used to predict the nitrate nitrogen and the soluble COD concentration in the secondary effluent. The error component is a time series model that utilizes the information about the past errors in order to predict the next one.

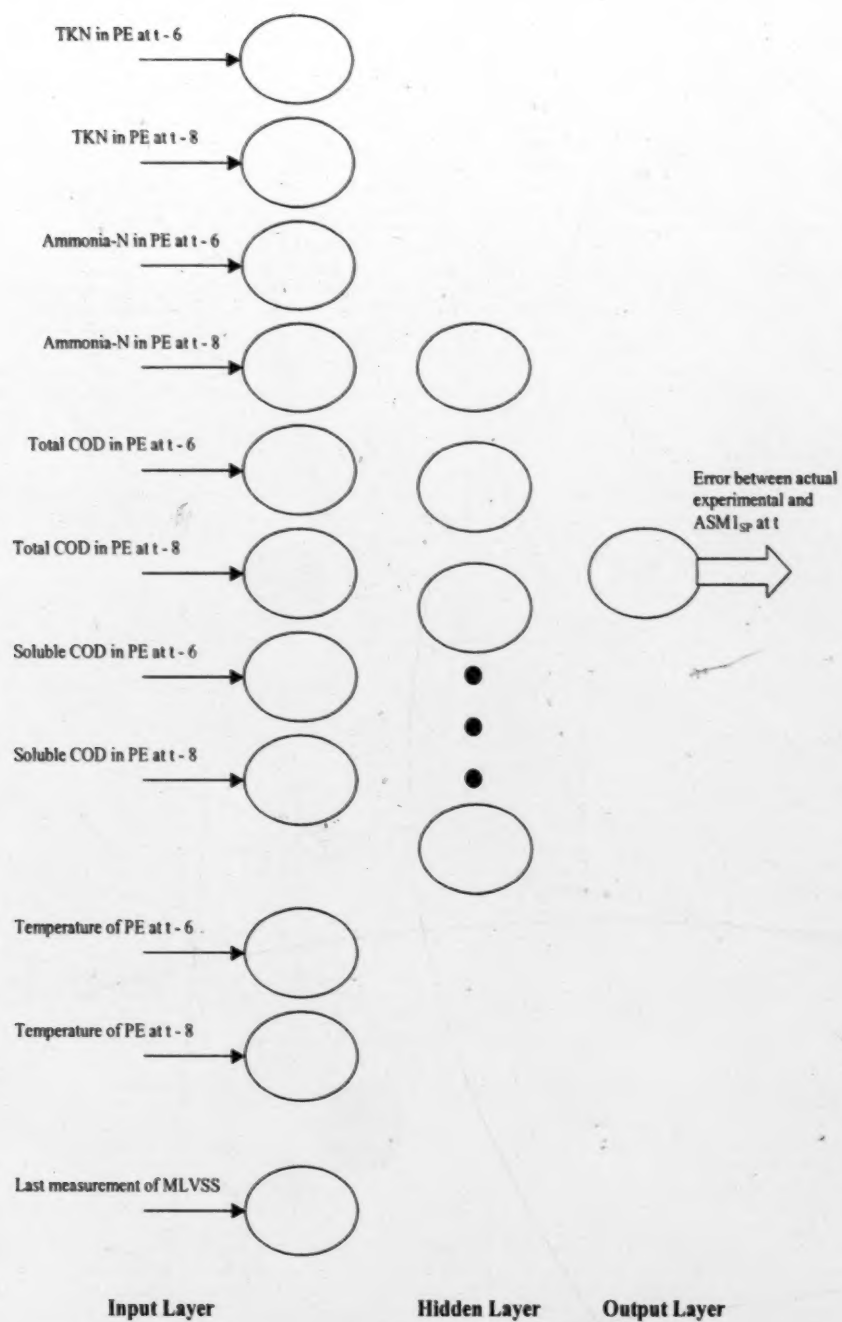


Figure 6.7 Artificial neural network error-predictor for ammonia nitrogen concentration in secondary effluent.

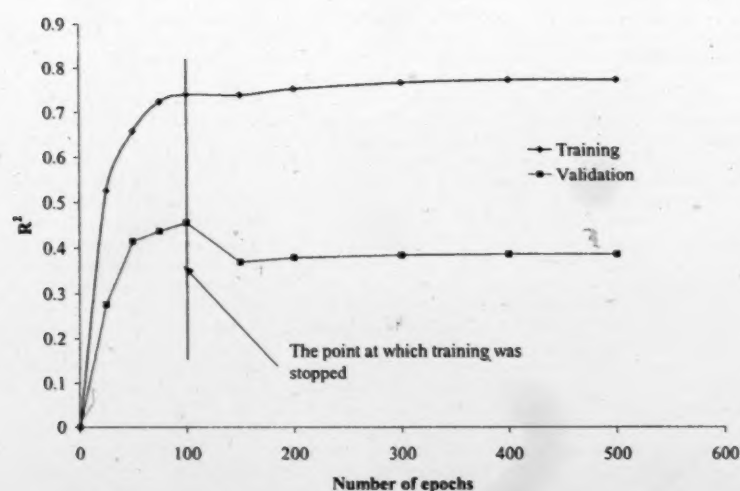


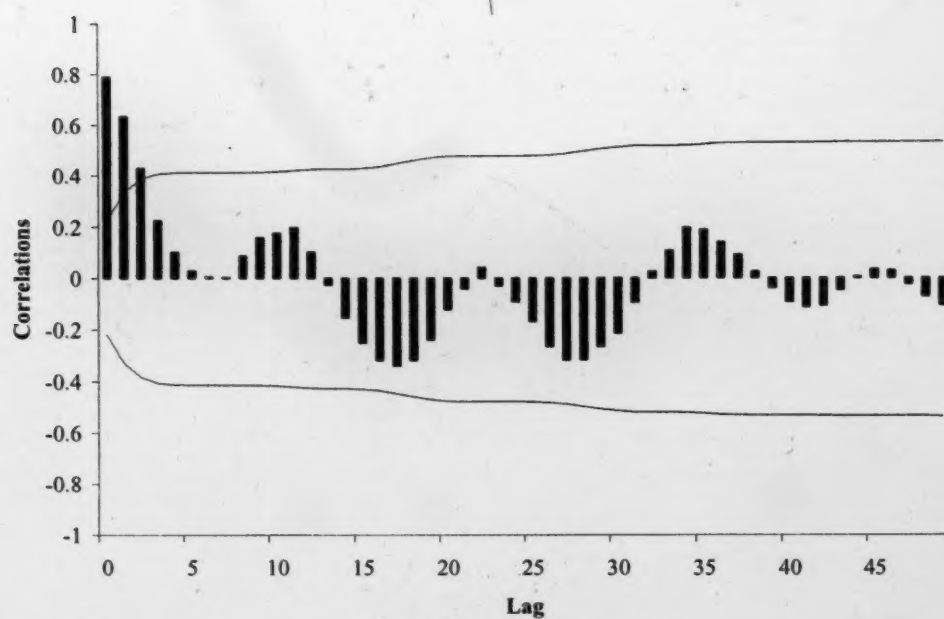
Figure 6.8 Monitoring the performance of the neural network error-predictor – Data of the first survey.

From the model identification step and after trying different candidate models, it was decided to use an autoregressive time series model of the first order as the error-predictor component for predicting the nitrate nitrogen concentration in the secondary effluent. Correlograms of the autocorrelation and the partial autocorrelation of the error series during the second sampling survey are shown in Figure 6.9. The autoregressive operator was estimated to be 0.851 with a standard error of 0.052 for the first survey while it was 0.798 with a standard error of 0.077 for the second survey.

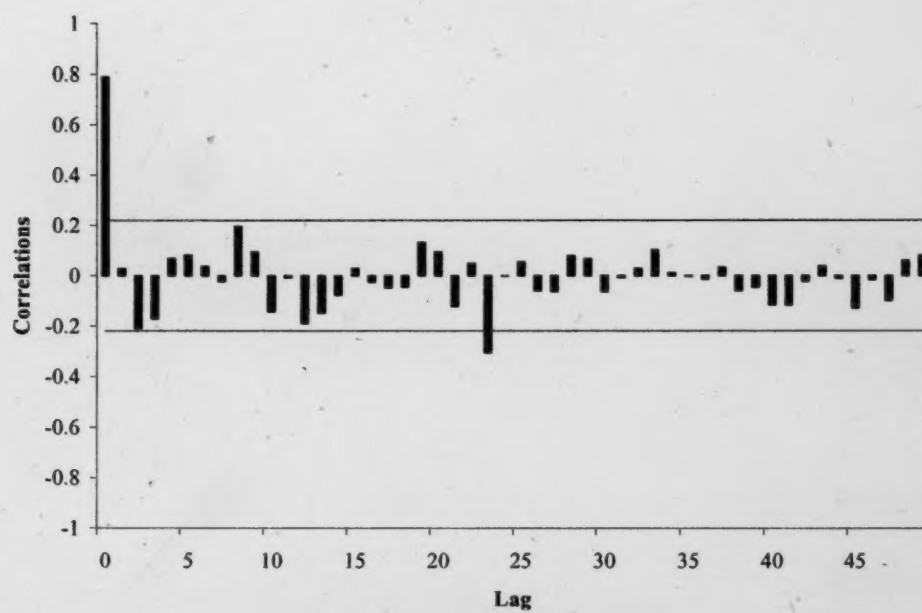
Table 6.1. The structure and parameter values for the time series models used as the error component in the hybrid model for predicting the soluble COD in the secondary effluent.

Structure and parameter values	First survey	Second survey
Model structure	ARMA(1,1)	AR(1)
Autoregressive parameters	AR(1) 0.841(0.081) ^a	AR(1) 0.527(0.112) ^a
Moving average parameters	MA(1) 0.377(0.141) ^a	N/A

^a Numbers in parentheses indicate the standard error



(a) Autocorrelations



(b) Partial autocorrelations

Figure 6.9. Correlograms of the autocorrelation and partial autocorrelation functions for the nitrate error series of the second survey. Solid lines represent the 95% confidence limits of two standard deviations.

For predicting the soluble COD concentration in the secondary effluent, a ARMA(1,1) model was used as the error-predictor component for the hybrid model for the data collected during the first survey, while a AR(1) error component was used in the hybrid model representing the data collected during the second survey. Table 6.1 lists the parameter values estimated for the models.

6.5.3 A Combined Transfer Function Noise Model for the Secondary Clarifier

The hybrid model that was described in the previous sections can only model the biological processes taking place in the activated sludge process, i.e., the physical settling characteristics of the mixed liquor in the secondary clarifier were not modelled by the hybrid model.

A simple transfer function noise model was used to predict the total suspended solids concentration in the secondary effluent at time t hours, using the flow of wastewater measured at the effluent weir and the total suspended solids in the mixed liquor entering the secondary clarifier both lagged by two hours, i.e., at time $t - 2$ hours. Different lagging periods within the range of theoretical detention times of the secondary clarifier were tried and the two-hour lag suited the data better.

Table 6.2 shows the structure of the models and the values estimated for the parameters. It is evident from Table 6.2 that some of the transfer function parameters estimated are lower than or almost equal to their standard error of estimate indicating that they might be insignificant. However, they were still included as both of the wastewater flow and the influent MLSS to the secondary clarifier are essential information for any kind of modelling for a secondary clarifier.

Table 6.2. The structure and parameter values for the models used to predict the total suspended solids concentration in the secondary effluent.

Survey no.	Output	Inputs	Model form and parameter estimates
1	Effluent TSS (Y_t)	Flow ($X_{1,t}$) MLSS ($X_{2,t}$)	$Y_t = \frac{\omega_{1,0}}{1} X_{1,t-1} + \frac{\omega_{2,0}}{1} X_{2,t-1} + \frac{(1-\theta_1 B)}{(1-\phi_1 B)} a_t$ $\omega_{1,0} = 0.060 (0.175)^a$ $\omega_{2,0} = 0.0028 (0.0029)^a$ $\phi_1 = 0.868 (0.07)^a$ $\theta_1 = 0.155 (0.136)^a$
2	Effluent TSS (Y_t)	Flow ($X_{1,t}$) MLSS ($X_{2,t}$)	$Y_t = \frac{\omega_{1,0}}{1} X_{1,t-1} + \frac{\omega_{2,0}}{1} X_{2,t-1} + \frac{1}{(1-\phi_1 B)} a_t$ $\omega_{1,0} = 0.028 (0.019)^a$ $\omega_{2,0} = 0.0011 (0.0005)^a$ $\phi_1 = 0.334 (0.170)^a$

^a Numbers in parentheses indicate the standard error

6.6 RESULTS AND DISCUSSION

As was mentioned previously, for each of the two sampling surveys that were conducted, data for the last two days were used in verifying the hybrid model while the rest of the data was used in training or fitting. Figure 6.10 shows the experimental data for the ammonia nitrogen concentration in the secondary effluent versus model predictions from which it is evident that the inclusion of the neural network error component substantially improved the predictions. Figure 6.10 shows that the cyclic nature of the actual data was well picked by the mechanistic component ASM1_{SP}. However, the peak values were in disagreement with the values predicted by ASM1_{SP} alone. The neural network component enabled the hybrid model to give peaks values that are in relatively good agreement with the actual experimental data.

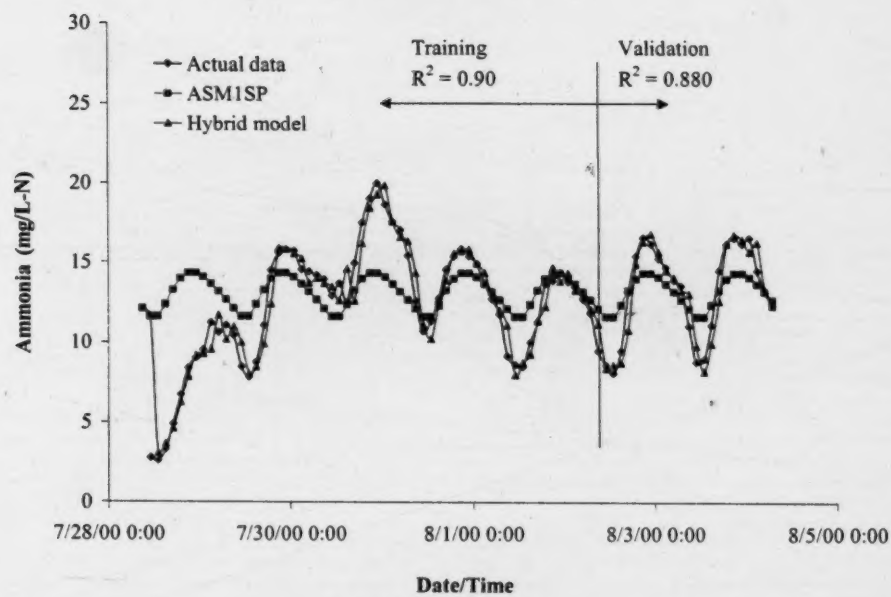
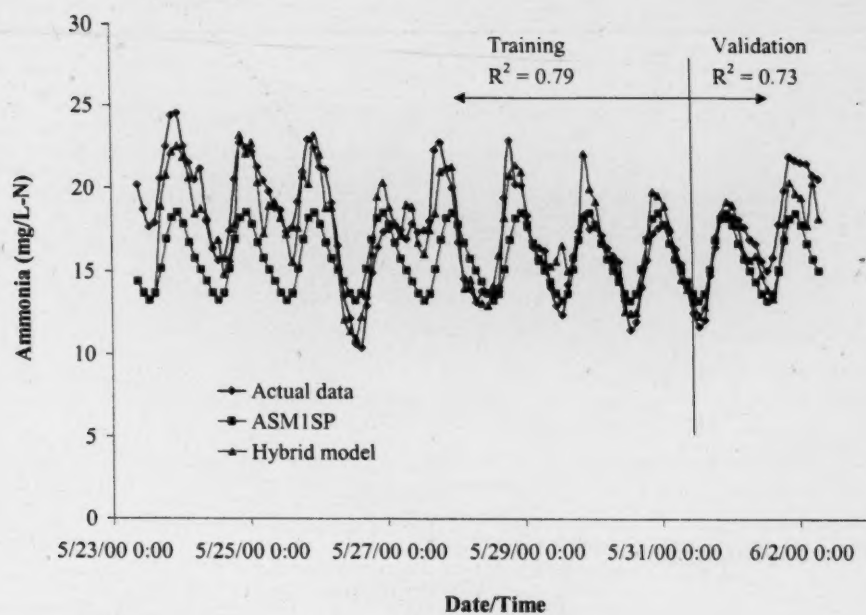
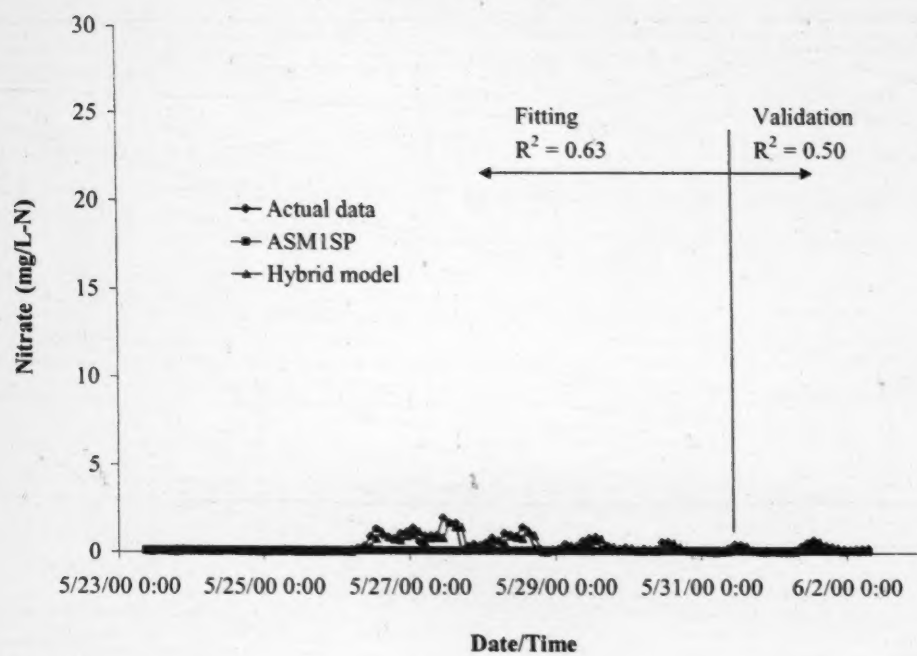
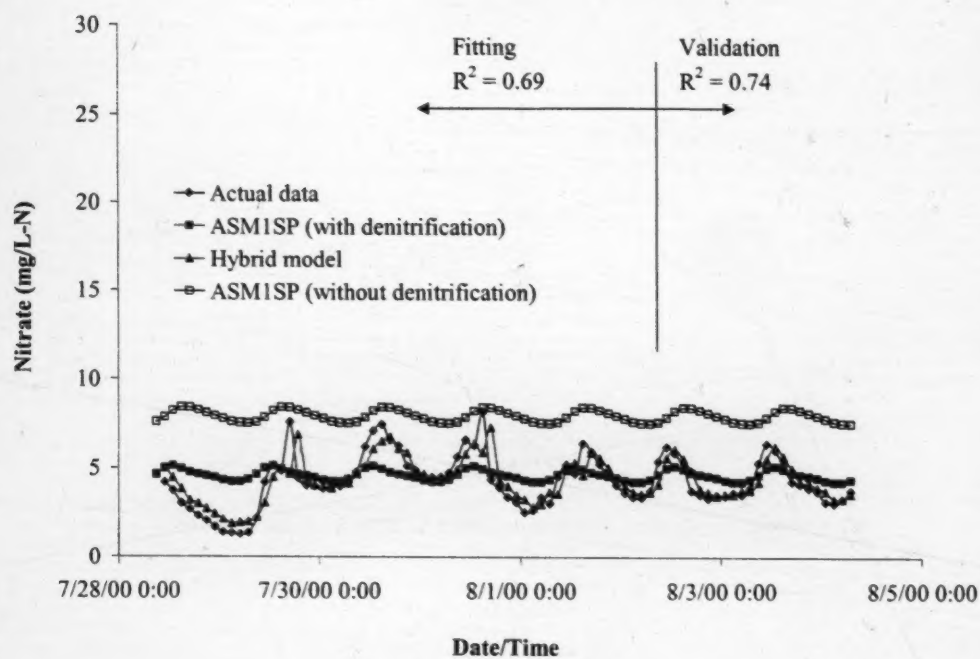


Figure 6.10. Actual data of the ammonia nitrogen concentration in the secondary effluent versus model predictions.

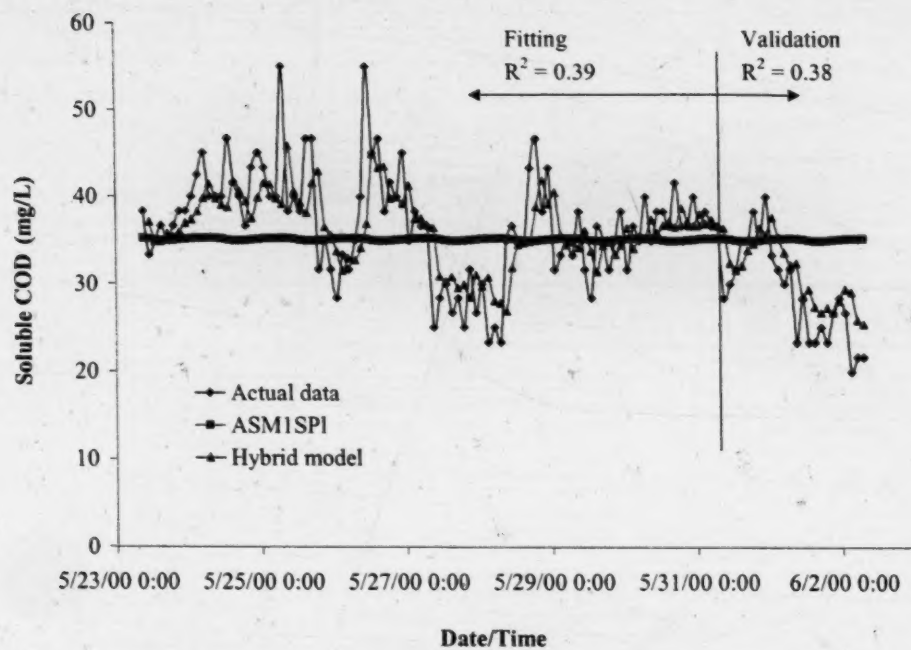


(a) First survey

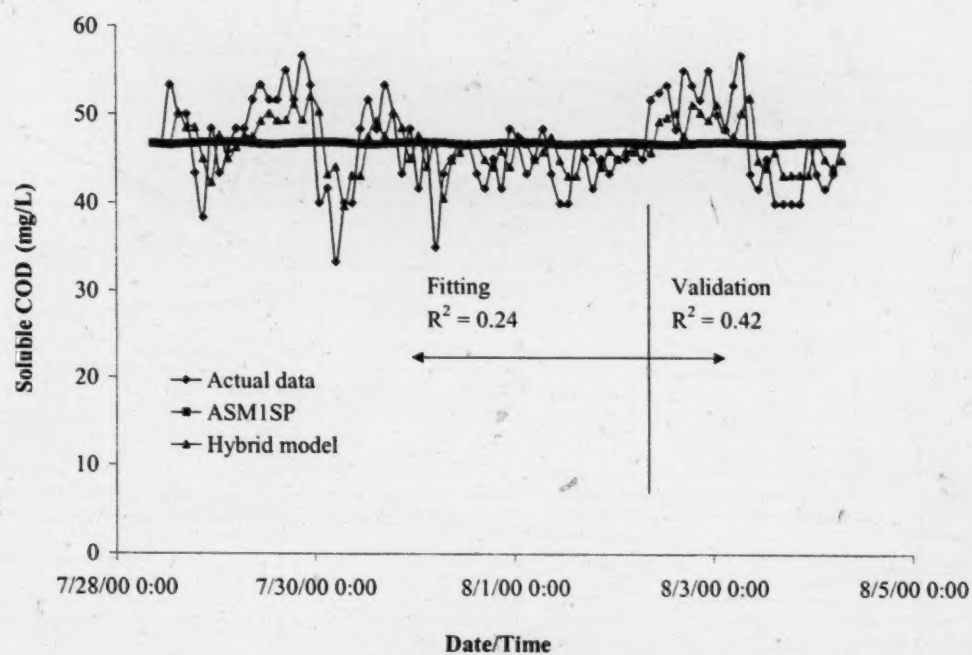


(b) Second survey

Figure 6.11. Actual data of the nitrate nitrogen concentration in the secondary effluent versus model predictions.

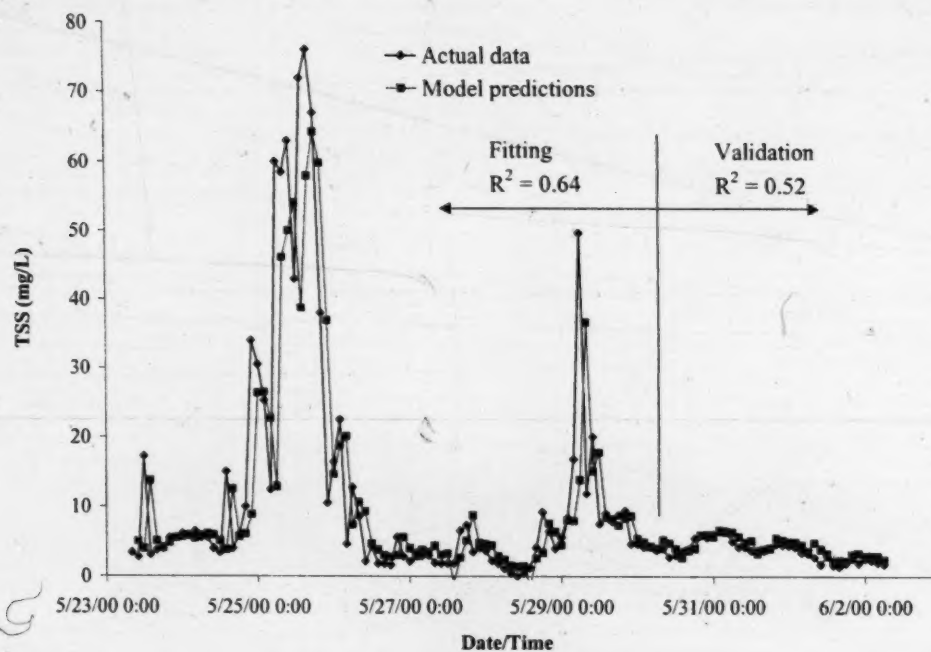


(a) First survey

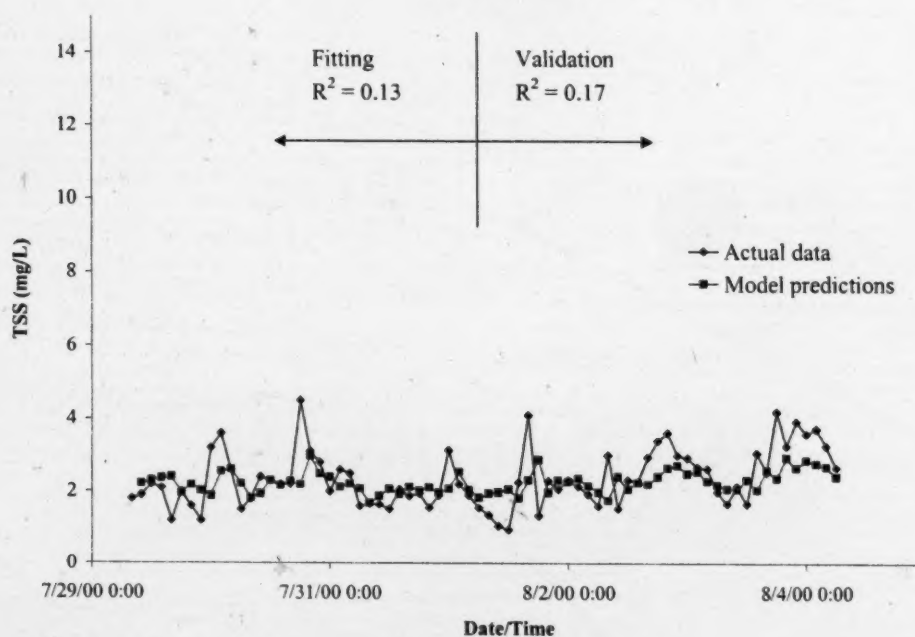


(b) Second survey

Figure 6.12. Actual data of the soluble COD concentration in the secondary effluent versus model predictions.



(a) First survey



(b) Second survey

Figure 6.13. Actual data of the total suspended solids concentration in the secondary effluent versus model predictions.

The experimental data for the nitrate nitrogen concentration in the secondary effluent versus model predictions are shown in Figure 6.11. The simulation results from ASM1_{SP} for the first sampling survey did not indicate any nitrification taking place in the aeration basin of the equivalent activated sludge tank. However, from the actual experimental data shown in Figure 6.11 (a), very low (< 1 mg/L) nitrate nitrogen concentrations were found in the secondary effluent samples collected. This discrepancy between the predictions by ASM1_{SP} and the actual data was accounted for by the time-series error-predictor component that was included in the hybrid model. The simulation by ASM1_{SP} during the second sampling survey indicated that partial nitrification took place in the aeration tank while denitrification took place in the sludge blanket of the secondary clarifier. As was mentioned before, the settling tests that were conducted by the plant operators during the period of the second survey clearly indicated that denitrification was taking place in the sludge blanket. In order to verify this finding, ASM1_{SP} simulations, with and without the inclusion of the sludge blanket of the secondary clarifier as part of the configuration, were conducted for the second sampling survey and the results are shown in Figure 6.11 (b). The modelling results shown in Figure 6.11 (b) clearly indicate that the denitrification in the sludge blanket had to be accounted for by the mechanistic component ASM1_{SP} in order to better describe the actual experimental data.

Figure 6.12 shows the experimental data for the soluble COD concentration in the secondary effluent versus model predictions. Although low values for the R^2 were attained by the hybrid model, the predictions by the mechanistic component ASM1_{SP} were substantially improved by the inclusion of the error-predictor component in the hybrid model. It is clear from Figure 6.12 that the deterministic model component, ASM1_{SP}, was not able to predict the dynamic variation in the secondary effluent soluble COD data. At the range of solids residence times (SRTs) encountered herein, almost all of the biodegradable soluble COD in the influent wastewater would be removed in the activated sludge process and only an extremely weak diurnal pattern could be identified in the unsteady state simulations provided by ASM1_{SP}. However not shown here, results of the steady state simulations that were conducted using ASM1_{SP} at similar values of the

SRT using wastewater characteristics, which are representative of the primary effluent at the GBWWTP clearly supported this finding. This discrepancy resulted from the fact that the inert soluble COD in the influent wastewater could not be measured on bi-hourly basis like the rest of the parameters measured. Therefore, an approximation had to be made in order to estimate an average value for the soluble inert COD in the influent wastewater. In addition, ASM1_{SP} does not model either the production of soluble microbial products by the microorganisms as they degrade the organic substrate in the influent to the reactor or the release of soluble cellular constituents through biomass lysis. Although little is known on the characteristics of these two types of soluble microbial products, they are thought to be biodegradable, although some at a very slow rate (Grady et al., 1999).

The experimental data for the total suspended solids concentration in the secondary effluent versus model predictions are shown in Figure 6.13. For the first sampling survey, reasonable agreement existed between the actual data and the predictions provided by the model. This was not the case for the second sampling survey. However, it should be noted that the total suspended solids concentrations during the second survey were in the lower range and averaged only 2.4 mg/L.

6.7 CONCLUSIONS

The sampling campaign that was conducted in the present study showed that partial nitrification took place in the activated sludge tank during the second survey, mainly due to the elevated temperatures encountered. Visual experimental observations during the settling tests and the unsteady state simulations by ASM1_{SP} indicated that denitrification took place in the secondary clarifier during the second survey. Therefore, the biological transformations occurring in the sludge blanket of the secondary clarifier had to be included as part of the activated sludge tank configuration.

Although ASM1_{SP} was able to give reasonable predictions for the ammonia nitrogen concentration in the secondary effluent, there were pronounced disagreement

between the actual data and the model's predictions for the soluble COD and nitrate nitrogen concentrations. Few simplifications were made in the model development of ASM1_{SP} as well as in the configuration that was used to depict the activated sludge process at the GBWWTP. In order to compensate for these simplifications, the hybrid model included a black box component that was used to predict the error between the actual data and the predictions by the deterministic component ASM1_{SP}. By including this error component in the hybrid model, the predictions by ASM1_{SP} were substantially improved.

For predicting the ammonia nitrogen concentration in the secondary effluent an artificial neural network model was used as the error component, while time series models were developed for the soluble COD and nitrate data using the Box-Jenkins methodology for model building. From a process-control point of view, the neural network error component used in the present study is better than the time series models that were used as the error-predictor components in the hybrid model. The reason is that the neural network model only utilizes information about the inputs to the system, however, the time series models use information about the past errors in order to predict the value for the next one. The problem with artificial neural networks that they need a substantial amount of training data in order for them to learn. Limited amount of data were available and therefore, in some cases artificial neural networks failed and time series models were used instead as the error-predictor component in the hybrid model.

In the present study, typical values that were available in the literature were used for the different kinetic and stoichiometric parameters in ASM1_{SP}. Although some of these parameters do not differ significantly from one domestic wastewater to another, and thus their values may be assumed, other parameters, such as Y_H and $B_{L,H}$, should be evaluated experimentally using treatability studies. By doing so, the deterministic component of the model will be improved and the load on the error-predictor component of the hybrid model will be reduced.

6.8 REFERENCES

1. Barker, P.S. and P.L. Dold, "General Model for Biological Nutrient Removal Activated-Sludge Systems: Model Presentation", *Wat. Env. Res.* 69: 969-984 (1997).
2. Box, G.E. and G.M. Jenkins, *Time Series Analysis: Forecasting and Control* (Oakland, CA, USA: Holden-Day, 1976).
3. Daigger, G.T. and D. Nolasco, "Evaluation and Design of Full-Scale Wastewater Treatment Plants Using Biological Process Models", *Wat. Sci. Tech.* 31(2): 245-255 (1995).
4. Eckenfelder, W.W., M.C. Goronszy, and T.P. Quirk, "The Activated Sludge Process: State of the Art", *Critical Reviews in Environmental Control* 15(2): 111-178 (1985).
5. Grady Jr., C.P.L., G.T. Daigger, and H.C. Lim, *Biological Wastewater Treatment*, 2nd ed. (NY, USA: Marcel Dekker, Inc., 1999).
6. Henze, M., C.P.L. Grady Jr., W. Gujer, G.v.R. Marais, and T. Matsuo, *Activated Sludge Model No. 1. IAWPRC Scientific and Technical Reports No. 1* (London, UK: IAWPRC, 1987).
7. Henze, M., W. Gujer, T. Mino, T. Matsuo, M.C. Wentzel, and G.v.R. Marais, *Activated Sludge Model No. 2. IAWQ Scientific and Technical Reports No. 3* (London, UK: IAWQ, 1995).
8. Henze, M., W. Gujer, T. Mino, and M. van Loosdrecht, *Activated Sludge Models ASM1, ASM2, ASM2d and ASM3. IAWQ Scientific and Technical Reports No. 9* (London, UK: IWA, 2000).
9. Lawrence, A.W. and P.L. McCarthy, "Unified Basis for Biological Treatment Design and Operation", *J. San. Eng., ASCE* 96(SA3): 757-778 (1970).
10. McKinney, R.E., "Mathematics of Complete-Mixing Activated Sludge", *J. San Eng., ASCE* 88(SA3): 87-113 (1962).
11. Monod, J., "The Growth of Bacterial Culture", *Annual Review of Microbiology* 3: 371-394 (1949).

12. Zhao, H., O.J. Hao, T.J. McAvoy, and C.H. Chang, "Modelling Nutrient Dynamics in Sequencing Batch Reactor", J. Env. Eng., ASCE 123(4): 311-319 (1997).

CHAPTER 7. CONCLUSIONS AND RECOMMENDATIONS

7.1 GENERAL OVERVIEW

Within the past few years, the topic of real time control (RTC) of wastewater treatment plants (WWTPs) has quickly become a promising and active area of research and the improvement of the computer-based infrastructure for WWTPs has been identified as an important priority. In order to obtain an effective RTC of a WWTP, the following dynamic processes have to be modelled: (1) input to the system, i.e., rainfall-runoff-sewer flow and quality models; and (2) system response to the input, i.e., models that have the ability to relate causes (inputs, controls) to effects (outputs, responses). In order to relate the inputs of an unsteady state system to its output, dynamic modelling techniques have to be utilized.

The main scope of this research project was to develop dynamic models that can describe the following: (1) the wastewater inflow entering the Gold Bar Wastewater Treatment Plant (GBWWTP); (2) the dynamic nature of the primary sedimentation process at the plant; and (3) the activated sludge process at the plant. The objective of the modelling effort was to develop models which have the potential of being used as part of a RTC system that can be designed to improve the performance of the different processes at the plant.

An artificial neural network (ANN) model was developed to provide short-term predictions of the wastewater inflow rate that enters the GBWWTP. The neural network model uses rainfall data, observed in the collection system discharging to the plant, as inputs. Among the 18 rain gauges that cover the major drainage basins of the city's sewerage system, 8 rain gauges were chosen to be used by the neural network. These gauges were selected because they are located in drainage areas that have high percentage of combined sewers. In order for the neural network model to learn the diurnal pattern of wastewater flow entering the plant on different days of the week during dry weather flow conditions, an index to represent the day-of-the-week and another index to represent the

hour-of-the-day were used as inputs to the neural network model. The available data set contained 114 rainfall events that were divided into three sets, namely, set #1, set #2, and set #3. Data set #1 was always used as part of the training data set. Data sets #2 and #3 were used as validation data sets and each contained 22 rainfall events. Each of the two validation data sets included rainfall events that covered a wide range of total amount of rainfall, maximum intensity, maximum flow at the entrance to the plant that was encountered during the rainfall event, and total inflow (amount of flow, conveyed to the plant during a rainfall event, that was above the normal dry weather flow). A moving window was used to present the rainfall data to the neural network model. The building process of the model was conducted in a systematic way. Different sizes of the moving window were tried. Only one hidden layer was used while different sizes of the hidden layer were tried. The 30-minute-in-advance wastewater flow was the output of the neural network. The objective was to predict the flow entering the plant during wet weather flow conditions, and as a result, the performance of the candidate models (with respect to training and validation) was judged only using the wet weather flow patterns.

The artificial neural network modelling approach was used to predict the dynamic response of a full-scale primary sedimentation tank at the GBWWTP. The neural network model consists of two separate networks, one uses flow and influent total suspended solids data in order to predict the effluent total suspended solids from the tank, and the other makes predictions of the effluent chemical oxygen demand using data of the flow and influent chemical oxygen demand as inputs. Two weeks of hourly sampling were conducted in order to collect a data set to be used in training and validating the networks. A systematic approach was used in the building process of the model. In order to present the input data records to the neural network, a moving window of past records was used. Two days worth of hourly data were used for validation while the rest of the data were used for training the models. The Box-Jenkins transfer function methodology was used in order to study how and to what extent effluent total suspended solids (TSS) and chemical oxygen demand (COD) are related to influent TSS, COD, and flow in the primary sedimentation tank. With the Box-Jenkins approach, stochastic and transfer function

components were combined to form a dynamic model and the relative importance of these two components was quantitatively assessed.

A hybrid model was used to predict the quality of effluent wastewater from a full-scale activated sludge process at the GBWWTP. The model consisted of two components: (1) mechanistic component, which is a simplified version of Activated Sludge Model No. 1 and was denoted by ASM1_{SP}; and (2) black-box component, which is an artificial neural network in some cases while a time series component in other cases. The function of the black-box component was to predict the error between the actual experimental data and the predictions from the mechanistic model component. By including the black-box component into the hybrid model, the simplifications that were applied to the ASM1 were compensated for. Because the objective was to calibrate and verify the model based on experimental data collected from full-scale operation, an extensive sampling campaign was conducted on one of the activated sludge tanks at the plant during the summer of 2000. Two sampling surveys were conducted as part of the campaign, one in May/June, during which the tank was operated in the plug flow mode, and the other in July/August during which the step-feed mode was used.

7.2 CONCLUSIONS

Based on the theoretical and experimental studies conducted in this research program, the following main conclusions can be drawn:

1. For almost all the candidate models that were able to converge, as the number of epochs that was used in training increased, the value of the R^2 for the training set increased and the maximum value was obtained when training was allowed for the maximum number of epochs that was used. On the other hand, the maximum value of the R^2 for the validation data set was always obtained at a much lower number of epochs (on one instance was as low as 30 epochs).
2. When the dimensionality of the network increased (either the size of the moving window or the size of the hidden layer is in the higher range) the maximum value of

the R^2 for the validation data set was obtained at low number of epochs and allowing training beyond this point hindered the ability of the network to generalize.

3. Although a network that is trained with a gradient descent algorithm, like the back-propagation, is not guaranteed to find a global minimum, even if it converges, the systematic approach of ANN model building that was used herein allowed the avoidance of local minimum.
4. For the ANN flow model, increasing the size of the moving window up to 8 hours had the effect of improving the performance of the network with respect to both training and validation. Expanding the moving window beyond 8 hours had very little effect on the performance.
5. The two indices of day-of-the-week and hour-of-the-day allowed the ANN flow model to learn the diurnal changes of the dry weather flow at the entrance to the plant as well as the lag that existed between the weekend patterns and the weekdays ones.
6. For both the training and validation data sets, the ANN flow model performed extremely well. When tested against the validation data that have not been seen by the model during the course of training, the model had excellent generalization ability and was able to predict fairly well the quantity of wastewater flow entering the plant for 44 different rainfall events that ranged widely in both the total amount of rainfall and the maximum intensity.
7. The ANN flow model was not confused by erroneous flow records in the training data set, which means that the model can be updated (or re-trained) online without visual examination of training data that are presented to the model as long as the overall quality of the data is good.
8. For the ANN models for the primary sedimentation tank, using a moving window of size 3 hours yielded the best results for the validation data set. Based on the results of the simple tracer studies conducted, the lag time between the effluent and the influent data used in modelling was set equal to one hour. When tested against the validation data set, the models were able to generalize well. The ANN model for

predicting the COD in the primary effluent gave better predictions than that of the model for TSS. The model predictions were better during wet weather conditions.

9. The sampling that was conducted on the primary sedimentation tank showed that during periods of rainfall events the TSS in the influent wastewater increased substantially due to the first flush effect while the COD values decreased due to the dilution effect of the storm water.
10. The ANN models for the primary sedimentation tank were superior than traditional non linear regression models.
11. For the combined transfer function noise models that were developed for the primary sedimentation tank, the addition of a transfer function component improved the prediction, and with the use of a transfer function model by itself (no noise component), the performance was worse than with a noise model by itself (no transfer function component). The identification of the delay parameter in the transfer function component clearly indicated the presence of short circuiting in the sedimentation tank. The two models used for the prediction of the effluent TSS (one for each week of sampling) had identical structures, however, they had different values for the estimated parameters, which means that updating such models could be done automatically without the need to change the structure of the model. The combined transfer function noise models gave overall better predictions than the ANN models that utilized the same data set, however, the ANN models have the advantage that they only use information about the input data (no previous values of the output variable are used as inputs).
12. The steady state simulations that were conducted for the activated sludge process @ 20°C showed that almost complete utilization of the readily biodegradable substrate would occur at an SRT as low as one day. Only very slight increase in the readily biodegradable substrate concentration in the secondary effluent would occur if the mode of operation is switched from plug flow to step-feed. The simulations also showed that nitrification does not take place in the aeration basin when the SRT is less than two days. However, at 20°C and an SRT of three days almost full nitrification could be established in the tank and increasing the SRT beyond three days would not have a pronounced effect on the percent nitrification. At an SRT of

three days there would be a slight increase in the ammonia-N concentration in the secondary effluent if the step-feed mode of operation is used instead of the conventional plug flow, however, beyond SRT of three days both the plug flow and the step-feed simulations gave virtually identical results.

13. The results of the two sampling surveys that were conducted on the activated sludge tank showed that although the tank was operated at an estimated SRT value of 2.1 days during the second survey while the estimated SRT value during the first survey was 2.3 days, partial nitrification took place in the aeration tank during the second survey. This was mainly due to the higher temperatures encountered during the second survey. Visual experimental observations during the settling tests conducted on the mixed liquor from the 4th pass of the aeration tank as well as the unsteady state simulations by ASM1_{SP} indicated that denitrification took place in the secondary clarifier during the second survey, and therefore, the biological transformations occurring in the sludge blanket of the secondary clarifier had to be included as part of the activated sludge tank configuration. Although denitrification took place in the sludge blanket of the secondary clarifier during the second survey, the effluent quality in terms of total suspended solids did not deteriorate. This was probably because nitrification did not occur to the full extent possible in the aeration tank, and therefore, the nitrate feed to the sludge blanket in the secondary clarifier did not produce nitrogen gas at a rate that was enough to cause the sludge flocs to rise to the effluent launders.
14. Although ASM1_{SP} was able to give reasonable predictions for the ammonia nitrogen concentration in the secondary effluent, there were significant disagreements between the actual data and the model's predictions for the soluble COD and nitrate nitrogen concentrations. The black-box error-predictor component of the hybrid model compensated for the few simplifications that were made in the model development of ASM1_{SP} as well as in the configuration that was used to depict the activated sludge process at the GBWWTP. For predicting the ammonia nitrogen concentration in the secondary effluent an artificial neural network model was used as the error-predictor component, while time series models were developed for the soluble COD and nitrate data using the Box-Jenkins methodology

for model building. By including this black-box component in the hybrid model, the predictions by ASM1_{SP} were substantially improved.

7.3 RECOMMENDATIONS

In this study, dynamic models were developed for predicting the wastewater inflow to the GBWWTP and for describing the dynamics of full-scale primary sedimentation and activated sludge tanks at the plant. Based on the achieved goals and the drawn conclusions, several recommendations have to be addressed in further future research:

1. Updating of the ANN flow model with recent data needs to be done. Using the same model structure, the new set of data that would be applied to the network for the purpose of re-estimating the weights. As the neural network model is presented with more scenarios of rainfall events, its prediction ability would improve.
2. The full-scale implementation of the ANN flow model is needed. In order to use the model as a part of an on-line control system, real-time predictions should be made during rainfall events. For this to be done, the rainfall data have to be downloaded from the remote locations of the rain gauges every 30 minute then prepared to be presented as inputs to the neural network, then the array of inputs can be processed through the trained network in order to predict the wastewater flow 30-minute-in-advance. This task has to be done every 30 minute in order to use the model on real-time basis. A simple computer program can accomplish this task.
3. More detailed tracer studies should be performed on the primary sedimentation tanks in order to verify the existence of short circuiting that was found from both the simple tracer studies that were performed and from the modelling effort that was conducted.
4. A more elaborated and detailed settling model for the secondary clarifier is needed. In order to do so, more data needs to be collected during periods of hydraulic and organic loading stresses.

5. In order to implement the type of quality models that were described herein as part of a real-time control scheme, a reliable on-line analyzer has to be installed at the influent and effluent points of both the primary sedimentation and the activated sludge tanks in order to collect the data needed. Pilot-scale testing of the different analyzers available for domestic wastewater should be done in order to select the most reliable one.
6. In the present study, typical values that were available in the literature were used for the different kinetic and stoichiometric parameters in ASM1_{SP}. Although some of these parameters do not differ significantly from one domestic wastewater to another, and thus their values may be assumed, other parameters, such as Y_H and $B_{L,H}$, should be evaluated experimentally using treatability studies. By doing so, the deterministic component of the model would be improved and the load on the error-predictor component of the hybrid model would be reduced.